Maximum Likelihood Estimation for Wishart processes

Aurélien Alfonsi[1], Ahmed Kebaier[2], and Clément Rey[1]

[1]CERMICS, École des Ponts, UPE, Projet MathRisk ENPC-INRIA-UMLV, Champs-sur-Marne, France
[2]Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539)

April 18, 2016

Abstract

In the last decade, there has been a growing interest to use Wishart processes for modelling, especially for financial applications. However, there are still few studies on the estimation of its parameters. Here, we study the Maximum Likelihood Estimator (MLE) in order to estimate the drift parameters of a Wishart process. We obtain precise convergence rates and limits for this estimator in the ergodic case and in some nonergodic cases. We check that the MLE achieves the optimal convergence rate in each case. Motivated by this study, we also present new results on the Laplace transform that extend the recent findings of Gnoatto and Grasselli [17] and are of independent interest.

Keywords : Wishart processes, Laplace transform, parameter inference, maximum likelihood, limit theorems, local asymptotic properties.
AMS MSC 2010: 62F12, 44A10, 60F05, 91B70.

1 Introduction and preliminary results

The goal of this paper is to study the maximum likelihood estimation of the parameters of Wishart processes. These processes have been introduced by Bru [7] and take values in the set of positive semidefinite matrices. Let $d \in \mathbb{N}^*$ denote the dimension, $\mathcal{M}_d$ be the set of real d-square matrices, $\mathcal{S}_d^+$ (resp. $\mathcal{S}_d^{+,*}$) be the subset of positive semidefinite (resp. definite) matrices, $\mathcal{S}_d$ (resp. $\mathcal{A}_d$) the subset of symmetric (resp. antisymmetric) matrices. Wishart processes are defined by the following SDE

$$\begin{cases} dX_t = \left[\alpha a^\top a + bX_t + X_t b^\top\right] dt + \sqrt{X_t} dW_t a + a^\top dW_t^\top \sqrt{X_t}, \quad t > 0 \\ X_0 = x \in \mathcal{S}_d^+, \end{cases} \tag{1}$$

where $\alpha \geqslant d-1$, $a \in \mathcal{M}_d$, $b \in \mathcal{M}_d$ and $(W_t)_{t\geqslant 0}$ denotes a d-square matrix made of independent Brownian motions. We recall that for $x \in \mathcal{S}_d^+$, $\sqrt{x}$ is the unique matrix in $\mathcal{S}_d^+$ such that

e-mails : alfonsi@cermics.enpc.fr, kebaier@math.univ-paris13.fr, reyc@cermics.enpc.fr. This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque and the Laboratories of Excellence Bézout http://bezout.univ-paris-est.fr/ and MME-DII http://labex-mme-dii.u-cergy.fr/ .

$\sqrt{x}^2 = x$. It is shown by Bru [7] and Cuchiero et al. [8] in a more general affine setting that the SDE (1) has a unique strong solution when $\alpha \geqslant d+1$ and a unique weak solution when $\alpha \geqslant d-1$. Besides, we have $X_t \in \mathcal{S}_d^{+,*}$ for any $t \geqslant 0$ when $x \in \mathcal{S}_d^{+,*}$ and $\alpha \geqslant d+1$. In this paper, we will denote by $WIS_d(x, \alpha, b, a)$ the law of $(X_t, t \geqslant 0)$ and $WIS_d(x, \alpha, b, a; t)$ the law of X_t. In dimension $d = 1$, Wishart processes are known as Cox-Ingersoll-Ross processes in the literature. It is worth recalling that the law of X only depends on a through $a^\top a$ since we have

$$WIS_d(x, \alpha, b, a) \underset{law}{=} WIS_d(x, \alpha, b, \sqrt{a^\top a}),$$

see e.g. equation (12) in [1]. Therefore, the parameters to estimate are α, b and $a^\top a$.

Wishart processes have been originally considered by Bru [6] to model some biological data. Recently, they have been widely used in financial models in order to describe the evolution of the dependence between assets. Namely, Gourieroux and Sufana [19] and Da Fonseca et al. [10] have proposed a stochastic volatility model for a basket of assets that assumes that the instantaneous covariance between the assets follows a Wishart process. This extends the well-known Heston model [21] to many assets. Wishart processes have also been used for interest rates models. Affine term structure models involving these processes have been proposed for example by Gourieroux and Sufana [20], Gnoatto [16] and Ahdida et al. [2]. For these models, the question of estimating the parameters of the underlying Wishart process may be important for practical purposes and should be possible thanks to the profusion of financial data. This issue has been considered by Da Fonseca et al. [9] for the model presented in [10]. However, there is no dedicated study on the Maximum Likelihood Estimator (MLE) for Wishart processes. For the Cox-Ingersoll-Ross process, the estimation of parameters has been studied earlier, motivated in particular by its use for interest rates (see Fournié and Talay [14]). Later on, the MLE has been studied by Overbeck [31] including some nonergodic cases, and more recently by Ben Alaya and Kebaier [4, 5]. This paper completes the literature by studying the MLE for Wishart processes.

In this paper, we will follow the theory developed in the books by Lipster and Shiryaev [27] and Kutoyants [23] and assume that we observe the full path $(X_t, t \in [0, T])$ up to time $T > 0$. This choice will be convenient from a mathematical point of view to study the convergence of the MLE. Of course, in practice it can be relevant to study precisely the estimation when we only observe the process on a discrete time-grid. This is left for further research, but we already observe in our numerical experiments that the discrete approximation of the MLE gives a satisfactory estimation of Wishart parameters (see Section 6). It is worth noticing that once we observe the path $(X_t, t \in [0, T])$, the parameter $a^\top a$ is known. In fact, we can calculate the quadratic covariation (see for example Lemma 2 in [1]) and get for $i, j, k, l \in \{1, \cdots, d\}$

$$\langle X_{i,j}, X_{k,l}\rangle_T = \int_0^T (a^\top a)_{j,l}(X_s)_{i,k} + (a^\top a)_{j,k}(X_s)_{i,l} + (a^\top a)_{i,l}(X_s)_{j,k} + (a^\top a)_{i,k}(X_s)_{j,l} ds. \tag{2}$$

This leads to

$$(a^\top a)_{i,i} = \frac{1}{4}\langle X_{i,i}\rangle_T \Big(\int_0^T (X_s)_{i,i} ds \Big)^{-1}, \tag{3}$$

$$(a^\top a)_{i,j} = \left(\frac{1}{2}\langle X_{i,j}, X_{i,i}\rangle_T - (a^\top a)_{i,i} \int_0^T (X_s)_{i,j} ds \right) \Big(\int_0^T (X_s)_{i,i} ds \Big)^{-1},$$

for $1 \leqslant i, j \leqslant d$ and $j \neq i$. We note that these quantities are well defined as soon as the path $(X_t, t \in [0,T])$ has a finite quadratic variation and is such that $X_t \in \mathcal{S}_d^{+,*}$ dt-a.e., which is satisfied by the paths of Wishart processes (see Proposition 4 in [7]). We will assume that $a^\top a \in \mathcal{S}_d^{+,*}$ and denote by $a \in \mathcal{M}_d$ an invertible matrix that matches the observed value of $a^\top a$: a can be for example the square root of $a^\top a$ or the Cholesky decomposition of $a^\top a$. Then, we know that $Y_t = (a^\top)^{-1} X_t a^{-1}$ follows the law $WIS_d((a^\top)^{-1} x a^{-1}, \alpha, (a^\top)^{-1} b a^\top, I_d)$, see e.g. equation (13) in [1]. It is therefore sufficient to focus on the estimation of the parameters α and b when $a = I_d$, which we consider now.

We first present the MLE of $\theta = (b, \alpha)$, and we denote by $\mathbb{P}_\theta$ the original probability measure under which X satisfies

$$dX_t = \left[\alpha I_d + bX_t + X_t b^\top\right] dt + \sqrt{X}_t dW_t + dW_t^\top \sqrt{X}_t. \tag{4}$$

When no confusion is possible, we also denote $\mathbb{P}$ this probability. We consider $\alpha_0 \geqslant d+1$ and set $\theta_0 = (\alpha_0, 0)$. We will assume for the joint estimation of α and b that

$$\alpha \geqslant d+1 \text{ and } x \in \mathcal{S}_d^{+,*}. \tag{5}$$

The latter assumption is not restrictive in practice since the condition $\alpha \geqslant d+1$ ensures that $X_t \in \mathcal{S}_d^{+,*}$ for any $t > 0$. Due to this assumption, we know by Theorem 4.1 in Mayerhofer [29] that

$$\frac{d\mathbb{P}_{\theta_0,T}}{d\mathbb{P}_{\theta,T}} := \exp\left(\int_0^T \mathrm{Tr}[H_s dW_s] - \frac{1}{2}\int_0^T \mathrm{Tr}[H_s H_s^\top] ds\right), \text{ with } H_t = \frac{\alpha_0 - \alpha}{2}(\sqrt{X_t})^{-1} - b\sqrt{X_t}$$

defines a probability measure under which $\tilde{W}_t = W_t - \int_0^t H_s^\top ds$ is a $d \times d$-Brownian motion, where $\mathbb{P}_{\theta,T}$ is the restriction of $\mathbb{P}_\theta$ to the σ-algebra $\sigma(W_s, s \in [0,T])$. We have

$$dX_t = \alpha_0 I_d dt + \sqrt{X}_t d\tilde{W}_t + d\tilde{W}_t^\top \sqrt{X}_t,$$

and the likelihood is then defined by (see Lipster and Shiryaev [27], Chapter 7)

$$L_T^{\theta,\theta_0} = \frac{1}{\mathbb{E}\left[\exp\left(\int_0^T \mathrm{Tr}[H_s dW_s] - \frac{1}{2}\int_0^T \mathrm{Tr}[H_s H_s^\top] ds\right) \Big| \mathcal{F}_T^X\right]}, \tag{6}$$

where $(\mathcal{F}_t^X)_{t \geqslant 0}$ denote the filtration generated by the process X.

Proposition 1.1. *For $X \in \mathcal{S}_d^{+,*}$, let $\mathcal{L}_X : \mathcal{S}_d \to \mathcal{S}_d$ be the linear application defined by $\mathcal{L}_X(Y) = XY + YX$. It is invertible, and the likelihood of $(X_t, t \in [0,T])$ is given by*

$$\begin{aligned} L_T^{\theta,\theta_0} = \exp\Big(&\frac{\alpha-\alpha_0}{4}\log\left(\frac{\det[X_T]}{\det[x]}\right) - \frac{\alpha-\alpha_0}{4}\big(\frac{\alpha+\alpha_0}{2} - 1 - d\big)\int_0^T \mathrm{Tr}[X_s^{-1}]ds - \frac{\alpha T}{2}\mathrm{Tr}[b] \\ &+ \frac{1}{2}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(bX_t + X_t b^\top\right) dX_t\right] - \frac{1}{4}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(bX_t + X_t b^\top\right)(bX_t + X_t b^\top)\right] dt\Big). \end{aligned} \tag{7}$$

Lemmas B.1 and B.2 states some properties of $\mathcal{L}_X$, and the proof of Proposition 1.1 is given in Appendix A. In particular, we see from this proof that $\frac{d\mathbb{P}_{\theta_0,T}}{d\mathbb{P}_{\theta,T}} \in \mathcal{F}_T^X$ if, and only if $b \in \mathcal{S}_d$,

in which case the likelihood has the following simpler form

$$L_T^{\theta,\theta_0} = \exp\Big(\frac{\alpha-\alpha_0}{4}\log\left(\frac{\det[X_T]}{\det[x]}\right) + \frac{\mathrm{Tr}[bX_T]-\mathrm{Tr}[bx]}{2} - \frac{1}{2}\int_0^T \mathrm{Tr}[b^2X_s]ds \\ - \frac{\alpha-\alpha_0}{4}\big(\frac{\alpha+\alpha_0}{2}-1-d\big)\int_0^T \mathrm{Tr}[X_s^{-1}]ds - \frac{\alpha T}{2}\mathrm{Tr}[b]\Big), \quad (8)$$

since $\mathcal{L}_{X_t}^{-1}(bX_t+X_tb)=b$.

Now, we want to maximize the likelihood and observe that the quantity in the exponential (7) is quadratic with respect to (b,α) and goes almost surely to $-\infty$ when $\|(b,\alpha)\|\to+\infty$. To do so, we first remark that $\mathrm{Tr}[b]=\mathrm{Tr}[\mathcal{L}_{X_t}^{-1}\left(bX_t+X_tb^\top\right)]$ by Lemma B.1. Then, Cauchy-Schwarz inequality yields to

$$\begin{aligned} |\mathrm{Tr}[\alpha b]| &= \left|\frac{1}{T}\int_0^T \mathrm{Tr}\left[\sqrt{2}\mathcal{L}_{X_s}^{-1}\left(bX_s+X_sb^\top\right)\sqrt{X_s}\frac{\alpha}{\sqrt{2}}\sqrt{X_s^{-1}}\right]ds\right| \\ &\leqslant \frac{1}{T}\int_0^T \mathrm{Tr}\left[(\mathcal{L}_{X_s}^{-1}\left(bX_s+X_sb^\top\right))^2X_s\right]ds + \frac{\alpha^2}{4}\frac{1}{T}\int_0^T \mathrm{Tr}\left[X_s^{-1}\right]ds \\ &= \frac{1}{2T}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_s}^{-1}\left(bX_s+X_sb^\top\right)(bX_s+X_sb^\top)\right]ds + \frac{\alpha^2}{4}\frac{1}{T}\int_0^T \mathrm{Tr}\left[X_s^{-1}\right]ds, \end{aligned} \quad (9)$$

and it is strict almost surely, which gives that the quadratic form in the exponential (7) is negative definite. There is thus a unique global maximum of (7) on $\mathbb{R}\times\mathcal{M}_d$. We know from Lemma B.2 that $\mathcal{L}_{X_s}^{-1}$ is self-adjoint, and we get with straightforward calculations that the MLE $\hat\theta_T=(\hat b_T,\hat\alpha_T)$ is characterized by the following equations:

$$\begin{cases} \frac{1}{4}\log\left(\frac{\det[X_T]}{\det[x]}\right) - \frac{\hat\alpha_T-1-d}{4}\int_0^T \mathrm{Tr}[X_s^{-1}]ds - \frac{T}{2}\mathrm{Tr}[\hat b_T]=0, \\ \int_0^T \mathcal{L}_{X_s}^{-1}(dX_s)X_s - \int_0^T \mathcal{L}_{X_s}^{-1}(\hat b_TX_s+X_s\hat b_T^\top)X_sds - \frac{\hat\alpha_TT}{2}I_d=0. \end{cases} \quad (10)$$

Unless in the ergodic case, we will not be able to obtain convergence results for this estimator. Instead, we will mostly work with the MLE estimator when b is known to be symmetric. This enables us to work with more tractable formulas, even if the calculations are already quite involved in case. Analyzing the general case would require development of further arguments. Besides, we can consider that Wishart processes with b symmetric already form an interesting family of processes that may be rich enough in many applications. When $b\in\mathcal{S}_d$, the unique global maximum $\hat\theta_T=(\hat b_T,\hat\alpha_T)$ of (8) on $\mathbb{R}\times\mathcal{S}_d$ is characterized by the following equations:

$$\begin{cases} \frac{1}{4}\log\left(\frac{\det[X_T]}{\det[x]}\right) - \frac{\hat\alpha_T-1-d}{4}\int_0^T \mathrm{Tr}[X_s^{-1}]ds - \frac{T}{2}\mathrm{Tr}[\hat b_T]=0, \\ \frac{X_T-x}{2} - \frac{1}{2}\int_0^T(\hat b_TX_s+X_s\hat b_T)ds - \frac{\hat\alpha_TT}{2}I_d=0. \end{cases} \quad (11)$$

To get more explicit formulas, we have to invert this linear system. For $X\in\mathcal{S}_d$ and $a\in\mathbb{R}$, we define the linear applications

$$\begin{aligned} \mathcal{L}_X:\ &\mathcal{S}_d\to\mathcal{S}_d & \text{and } \mathcal{L}_{X,a}:\ &\mathcal{S}_d\to\mathcal{S}_d \\ &Y\mapsto YX+XY & &Y\mapsto YX+XY-2a\,\mathrm{Tr}[Y]I_d. \end{aligned} \quad (12)$$

We introduce the following shorthand notation

$$R_T:=\int_0^T X_sds,\quad Q_T:=\left(\int_0^T \mathrm{Tr}[X_s^{-1}]ds\right)^{-1},\quad Z_T:=\log\left(\frac{\det[X_T]}{\det[x]}\right), \quad (13)$$

and note that Q_T and Z_T are defined only for $\alpha \geqslant d+1$ while R_T is defined for $\alpha \geqslant d-1$ and belongs almost surely to $\mathcal{S}_d^{+,*}$.[1] By using the convexity property of the inverse, see e.g. Mond and Pecaric [30], we have when $\alpha \geqslant d+1$

$$\mathrm{Tr}\left[\left(\frac{R_T}{T}\right)^{-1}\right] < \frac{Q_T^{-1}}{T},\ a.s. \tag{14}$$

We get $\hat{\alpha}_T = 1+d+Q_T\big(Z_T - 2T\,\mathrm{Tr}[\hat{b}_T]\big)$ and $\mathcal{L}_{R_T,T^2Q_T}(\hat{b}_T) = X_T - x - T\left(Q_T Z_T + 1 + d\right) I_d$. By (14) and Lemma B.1, the latter equation can be inverted, which leads to

$$\left\{\begin{array}{lcl} \hat{\alpha}_T & = & 1+d+Q_T\Big(Z_T - 2T\,\mathrm{Tr}\left[\mathcal{L}^{-1}_{R_T,T^2Q_T}\left(X_T - x - T\left[Q_T Z_T + 1 + d\right] I_d\right)\right]\Big) \\ \hat{b}_T & = & \mathcal{L}^{-1}_{R_T,T^2Q_T}\left(X_T - x - T\left[Q_T Z_T + 1 + d\right] I_d\right). \end{array}\right. \tag{15}$$

The estimator of α when $\alpha \in [d-1, d+1)$ given by the MLE is no longer well defined. The same thing already occurs in dimension $d=1$ for the CIR process, see Ben Alaya and Kebaier [4]. However, it is still possible to estimate the parameter $b \in \mathcal{M}_d$ when $\alpha \geqslant d-1$ is known. In this case, we denote $\theta = (b, \alpha)$ and $\theta_0 = (0, \alpha)$ and get by repeating the same arguments that

$$\begin{aligned} L_T^{\theta,\theta_0} = \exp\Big(&\frac{1}{2}\int_0^T \mathrm{Tr}\left[\mathcal{L}^{-1}_{X_t}\left(bX_t + X_t b^\top\right) dX_t\right] \\ &- \frac{1}{4}\int_0^T \mathrm{Tr}\left[\mathcal{L}^{-1}_{X_t}\left(bX_t + X_t b^\top\right)(bX_t + X_t b^\top)\right] dt - \frac{\alpha T}{2}\mathrm{Tr}[b]\Big), \end{aligned}$$

and the MLE is characterized by

$$\int_0^T \mathcal{L}^{-1}_{X_s}(dX_s)X_s - \int_0^T \mathcal{L}^{-1}_{X_s}(\hat{b}_T X_s + X_s \hat{b}_T^\top)X_s ds - \frac{\alpha T}{2} I_d = 0. \tag{16}$$

When b is known a priori to be symmetric, the likelihood and the MLE are then given by

$$L_T^{\theta,\theta_0} = \exp\Big(\frac{\mathrm{Tr}[bX_T] - \mathrm{Tr}[bx]}{2} - \frac{1}{2}\int_0^T \mathrm{Tr}[b^2 X_s] ds - \frac{\alpha T}{2}\mathrm{Tr}[b]\Big), \tag{17}$$

$$\hat{b}_T = \mathcal{L}^{-1}_{R_T}\left(X_T - x - \alpha T I_d\right). \tag{18}$$

The goal of the paper is to study the convergence of the MLE under the original probability $\mathbb{P}_\theta$. To do so, we first consider the case where the Wishart process is ergodic. By Lemma C.1, this holds if $-(b+b^\top) \in \mathcal{S}_d^{+,*}$ when $b \in \mathcal{M}_d$, and the ergodicity is equivalent to $-b \in \mathcal{S}_d^{+,*}$ when $b \in \mathcal{S}_d$. Then, we can use Birkhoff's ergodic theorem to determine the convergence of the MLE. Section 2 presents these results for (15) when $\alpha \geqslant d+1$, for (7) when $\alpha > d+1$ and for both (18) and (16) when $\alpha \geqslant d-1$. Section 3 studies the convergence of the MLE in some nonergodic cases, namely when $b = \lambda_0 I_d$ with $\lambda_0 \geqslant 0$ and when b is known to be symmetric. More precisely, when $b=0$, we obtain convergence results for (15) when $\alpha \geqslant d+1$ and for (18) when $\alpha \geqslant d-1$. When $\lambda_0 > 0$, we only obtain convergence results

[1] This is obvious when $\alpha > d-1$ since $X_t \in \mathcal{S}_d^{+,*}$ a.s. by Proposition 4 in [7]. For $\alpha = d-1$, we would have by contradiction the existence of $v_T \in \mathcal{F}_T^X$ such that $\forall t \in [0,T], v_T^\top X_t v_T = 0$. This is clearly not possible by using the connection with matrix-valued Ornstein-Uhlenbeck in this case, see eq. (5.7) in [7].

for (18) when $\alpha \geqslant d-1$. In all these cases, we analyse the convergence by the mean of Laplace transforms. Though limited to some nonergodic cases, we however recover and extend the recent convergence results obtained by Ben Alaya and Kebaier for the one-dimensional CIR process [4, 5]. In Section 4, we check that the MLE achieves the optimal rate of convergence in the different cases by proving local asymptotic properties. Last, we study in Section 5 the Laplace transform of (X_T, R_T). This study can be of independent interest and improves the recent results of Gnoatto and Grasselli [17].

2 Statistical Inference of the Wishart process: the ergodic case

When $-(b+b^\top) \in \mathcal{S}_d^{+,*}$, the Wishart process X_t converges in law when $t \to +\infty$ to the stationary law $X_\infty \sim WIS_d(0, \alpha, 0, \sqrt{2q_\infty}; 1/2)$ with $q_\infty = \int_0^\infty e^{sb}e^{sb^\top} ds$ for any starting point $x \in \mathcal{S}_d^+$ by Lemma C.1. Therefore this is the unique stationary law which is thus extremal, and we know by Stroock ([35], Theorem 7.4.8) that it is then ergodic, see also Pagès [32], Annex A. We introduce the following quantity

$$\overline{R}_\infty := \mathbb{E}_\theta(X_\infty).$$

From the ergodic Birkhoff's theorem, we have

$$\frac{R_T}{T} \xrightarrow{a.s.} \overline{R}_\infty, \qquad \text{as } T \to +\infty. \tag{19}$$

Besides, when $\alpha \geqslant d+1$, $\overline{Q}_\infty = \frac{1}{\mathbb{E}_\theta(\mathrm{Tr}[X_\infty^{-1}])}$ is finite and satisfies

$$\overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}] < 1, \tag{20}$$

due to the convexity property of the inverse, see e.g. Mond and Pecaric [30]. Again, the ergodic Birkhoff's theorem gives

$$TQ_T \xrightarrow{a.s.} \overline{Q}_\infty = \frac{1}{\mathbb{E}_\theta(\mathrm{Tr}[X_\infty^{-1}])}, \qquad \text{as } T \to +\infty. \tag{21}$$

This section is organized as follows. First, we study the MLE (15) when b is known to be symmetric in the cases $\alpha > d+1$ and $\alpha = d+1$. Then, we focus on the MLE (10) when $b \in \mathcal{M}_d$ and $\alpha > d+1$. The analysis follows the same steps and reuses some calculations made in the symmetric case. Last, we study the convergence of the MLE when $\alpha \geqslant d-1$ is known, in both symmetric and general cases.

2.1 The global MLE estimator of $\theta = (b, \alpha)$ when b is known to be symmetric

When $b \in \mathcal{S}_d$, the ergodicity is by Lemma (C.1) equivalent to $-b \in \mathcal{S}_d^{+,*}$, which we assume in this subsection. We have $X_\infty \sim WIS_d(0, \alpha, 0, \sqrt{-b^{-1}}; 1/2)$ and it is easy to get from (4) that $\alpha I_d + b\overline{R}_\infty + \overline{R}_\infty b = 0$, which gives $\overline{R}_\infty = -\frac{\alpha}{2}b^{-1} \in \mathcal{S}_d^{+,*}$. We will also show in the proof of Theorem 2.1 that

$$\overline{Q}_\infty := \frac{\alpha - (1+d)}{2\operatorname{Tr}[-b]}. \tag{22}$$

We consider the convergence of the MLE given by (15) when $\alpha \geqslant d+1$. We introduce the following martingales:

$$M_t := \int_0^t \sqrt{X_s}dW_s + \int_0^t dW_s^\top \sqrt{X}_s, \tag{23}$$

$$N_t := \int_0^t \mathrm{Tr}[(\sqrt{X_s})^{-1}dW_s]. \tag{24}$$

We use the dynamics of $(X_t)_{t\geqslant 0}$ under $\mathbb{P}_\theta$ and Itô's formula for $(Z_t)_{t\geqslant 0}$ (see e.g. Bru [7], equation (2.6)) to get on the one hand

$$X_T = x + \alpha T I_d + \mathcal{L}_{R_T}(b) + M_T, \qquad Z_T = (\alpha - 1 - d)Q_T^{-1} + 2\,\mathrm{Tr}[b]T + 2N_T. \tag{25}$$

On the other hand, we obtain from (11) and (13) that $X_T = x + \hat{\alpha}_T T I_d + \mathcal{L}_{R_T}(\hat{b}_T)$ and $Z_T = (\hat{\alpha}_T - 1 - d)Q_T^{-1} + 2T\,\mathrm{Tr}[\hat{b}_T]$, which yields to

$$\left\{ \begin{array}{rcl} \hat{\alpha}_T - \alpha &=& 2TQ_T\,\mathrm{Tr}[b - \hat{b}_T] + 2Q_T N_T \\ \mathcal{L}_{R_T}(\hat{b}_T - b) &=& (\alpha - \hat{\alpha}_T)TI_d + M_T = 2T^2 Q_T\,\mathrm{Tr}[\hat{b}_T - b]I_d + M_T - 2TQ_T N_T I_d. \end{array} \right. \tag{26}$$

Theorem 2.1. *Assume that $-b \in S_d^{+,*}$ and $\alpha > d+1$. Under $\mathbb{P}_\theta$, $\left(\sqrt{T}(\hat{b}_T - b, \hat{\alpha}_T - \alpha)\right)$ converges in law when $T \to +\infty$ to the centered Gaussian vector $(\mathbf{G}, H)$ that takes values in $\mathcal{S}_d \times \mathbb{R}$ and has the following Laplace transform: for $c, \lambda \in \mathcal{S}_d \times \mathbb{R}$,*

$$\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}] + \lambda H\right)\right] = \exp\left(\frac{2\overline{Q}_\infty \lambda^2}{1 - \overline{Q}_\infty\,\mathrm{Tr}[\overline{R}_\infty^{-1}]} - \frac{2\overline{Q}_\infty \lambda}{1 - \overline{Q}_\infty\,\mathrm{Tr}[\overline{R}_\infty^{-1}]}\,\mathrm{Tr}[c\overline{R}_\infty^{-1}] + \mathrm{Tr}[c\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)] \right).$$

Proof. By (14) and Lemma B.1, we can rewrite the system (26) as follows

$$\left\{ \begin{array}{rcl} \sqrt{T}(\hat{\alpha}_T - \alpha) &=& 2TQ_T \dfrac{N_T}{\sqrt{T}} - 2TQ_T\,\mathrm{Tr}\left[\mathcal{L}^{-1}_{\frac{R_T}{T}, TQ_T}\left(\dfrac{M_T}{\sqrt{T}} - 2TQ_T I_d \dfrac{N_T}{\sqrt{T}}\right)\right] \\ \sqrt{T}(\hat{b}_T - b) &=& \mathcal{L}^{-1}_{\frac{R_T}{T}, TQ_T}\left(\dfrac{M_T}{\sqrt{T}} - 2TQ_T I_d \dfrac{N_T}{\sqrt{T}}\right). \end{array} \right.$$

Note that, for $i,j,k,l \in \{1, \dots, d\}$ we have

$$\begin{array}{rcl} \langle M_{i,j}, M_{k,l}\rangle_t &=& \left[\delta_{jl}(R_t)_{i,k} + \delta_{jk}(R_t)_{i,l} + \delta_{il}(R_t)_{j,k} + \delta_{ik}(R_t)_{j,l}\right], \\ \langle M_{i,j}, N\rangle_t &=& 2t\delta_{ij} \quad \text{and} \quad \langle N\rangle_t = Q_t^{-1}, \end{array} \tag{27}$$

where δ_{ij} stands for the Kronecker symbol.

So, it follows from the central limit theorem for martingales (see e.g., Kutoyants [23], Proposition 1.21), that $(\frac{M_T}{\sqrt{T}}, \frac{N_T}{\sqrt{T}})$ converges in law under $\mathbb{P}_\theta$ towards a centered Gaussian vector $(\tilde{\mathbf{G}}, \tilde{H})$ taking values in $\mathcal{S}_d \times \mathbb{R}$ such that

$$\begin{array}{c} \mathbb{E}_\theta(\tilde{\mathbf{G}}_{i,j}\tilde{\mathbf{G}}_{k,l}) = \left[\delta_{jl}(\overline{R}_\infty)_{i,k} + \delta_{jk}(\overline{R}_\infty)_{i,l} + \delta_{il}(\overline{R}_\infty)_{j,k} + \delta_{ik}(\overline{R}_\infty)_{j,l}\right], \\ \mathbb{E}_\theta(\tilde{\mathbf{G}}_{i,j}\tilde{H}) = 2\delta_{i,j} \text{ and } \mathbb{E}_\theta(\tilde{H}^2) = \overline{Q}_\infty^{-1}. \end{array} \tag{28}$$

From (25) and (21), we obtain (22). From Lemma B.1, the function $(X, Y, a) \mapsto \mathcal{L}^{-1}_{X,a}(Y)$ is continuous, and we get by Slutsky's theorem that $(\sqrt{T}(\hat{b}_T - b), \sqrt{T}(\hat{\alpha}_T - \alpha))$ converges in law to the Gaussian vector

$$(\mathbf{G}, H) = \left(\mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}\left(\tilde{\mathbf{G}} - 2\overline{Q}_\infty \tilde{H} I_d\right), 2\overline{Q}_\infty\left(\tilde{H} - \mathrm{Tr}\left[\mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}\left(\tilde{\mathbf{G}} - 2\overline{Q}_\infty \tilde{H} I_d\right)\right]\right)\right).$$

We are interested to calculate the Laplace transform of this law. First, we calculate the Laplace transform of $(\tilde{\mathbf{G}}, \tilde{H})$:

$$\forall c \in \mathcal{S}_d, \lambda \in \mathbb{R}, \mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\tilde{\mathbf{G}}] + \lambda\tilde{H}\right)\right] = \exp\left(\frac{1}{2}\left(\lambda^2\overline{Q}_\infty^{-1} + 4\lambda\,\mathrm{Tr}[c] + 4\,\mathrm{Tr}[c^2\overline{R}_\infty]\right)\right). \tag{29}$$

We want to calculate for $c \in \mathcal{S}_d$ and $\lambda \in \mathbb{R}$,

$$\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}] + \lambda H\right)\right] = \mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[(c - 2\lambda\overline{Q}_\infty I_d)\mathbf{G}] + 2\lambda\overline{Q}_\infty\tilde{H}\right)\right].$$

Due to (20) and Lemma B.1, we can introduce $\tilde{c} = \mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}(c - 2\lambda\overline{Q}_\infty I_d)$. We have

$$\overline{R}_\infty\tilde{c} + \tilde{c}\overline{R}_\infty - 2\overline{Q}_\infty\,\mathrm{Tr}[\tilde{c}]I_d = c - 2\lambda\overline{Q}_\infty I_d,$$

and thus

$$\begin{aligned}\mathrm{Tr}[(c - 2\lambda\overline{Q}_\infty I_d)\mathbf{G}] &= \mathrm{Tr}[(\overline{R}_\infty\tilde{c} + \tilde{c}\overline{R}_\infty - 2\overline{Q}_\infty\,\mathrm{Tr}[\tilde{c}]I_d)\mathbf{G}] \\ &= \mathrm{Tr}[\tilde{c}(\overline{R}_\infty\mathbf{G} + \mathbf{G}\overline{R}_\infty - 2\overline{Q}_\infty\,\mathrm{Tr}[\mathbf{G}]I_d)] = \mathrm{Tr}[\tilde{c}(\tilde{\mathbf{G}} - 2\overline{Q}_\infty\tilde{H}I_d)].\end{aligned}$$

We therefore obtain from (29)

$$\begin{aligned}&\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}] + \lambda H\right)\right] \\ &= \mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[\tilde{c}(\tilde{\mathbf{G}} - 2\overline{Q}_\infty\tilde{H}I_d)] + 2\lambda\overline{Q}_\infty\tilde{H}\right)\right] \\ &= \mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[\tilde{c}\tilde{\mathbf{G}}] + 2\overline{Q}_\infty(\lambda - \mathrm{Tr}[\tilde{c}])\tilde{H}\right)\right] \\ &= \exp\left(2\left\{(\lambda - \mathrm{Tr}[\tilde{c}])^2\overline{Q}_\infty + 2(\lambda - \mathrm{Tr}[\tilde{c}])\,\mathrm{Tr}[\tilde{c}]\overline{Q}_\infty + \mathrm{Tr}[\tilde{c}^2\overline{R}_\infty]\right\}\right).\end{aligned}$$

Since $2\,\mathrm{Tr}[\tilde{c}^2\overline{R}_\infty] = \mathrm{Tr}[\tilde{c}(\tilde{c}\overline{R}_\infty + \overline{R}_\infty\tilde{c})] = \mathrm{Tr}[\tilde{c}c] + 2\overline{Q}_\infty(\mathrm{Tr}[\tilde{c}] - \lambda)\,\mathrm{Tr}[\tilde{c}]$, we get

$$\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}] + \lambda H\right)\right] = \exp\left(2\lambda(\lambda - \mathrm{Tr}[\tilde{c}])\overline{Q}_\infty + \mathrm{Tr}[\tilde{c}c]\right).$$

We now use that $\mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}(I_d) = \frac{1}{2(1-\overline{Q}_\infty\,\mathrm{Tr}[\overline{R}_\infty^{-1}])}\overline{R}_\infty^{-1}$ to get $\tilde{c} = \mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}(c) - \lambda\frac{\overline{Q}_\infty\overline{R}_\infty^{-1}}{1-\overline{Q}_\infty\,\mathrm{Tr}[\overline{R}_\infty^{-1}]}$. Since we have $\mathrm{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty,\overline{Q}_\infty}(c)] = \frac{\mathrm{Tr}[\overline{R}_\infty^{-1}c]}{2(1-\overline{Q}_\infty\,\mathrm{Tr}[\overline{R}_\infty^{-1}])}$ by Lemma B.1, this yields to the claimed result. □

When $\alpha = d+1$, the rate of convergence of the MLE of α is even better as stated by the following theorem.

Theorem 2.2. *Assume $-b \in S_d^{+,*}$ and $\alpha = d+1$. Then, under $\mathbb{P}_\theta$, $\left(\sqrt{T}(\hat{b}_T - b), T(\hat{\alpha}_T - \alpha)\right)$ converges in law when $T \to +\infty$ to $\left(\mathbf{G}, -2\tau^{-1}_{-\mathrm{Tr}[b]}\,\mathrm{Tr}[b]\right)$, where $\tau_a = \inf\{t \geq 0,\ B_t = a\}$ with $(B_t)_{t\geq 0}$ a given one-dimensional standard Brownian motion and $\mathbf{G}$ is a Gaussian vector independent of B such that $\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}]\right)\right] = \exp\left(\mathrm{Tr}[c\mathcal{L}^{-1}_{\overline{R}_\infty}(c)]\right)$, $c \in \mathcal{S}_d$.*

Proof. By (14) and Lemma B.1, we can rewrite the system (26) as follows

$$\begin{cases} T(\hat{\alpha}_T - \alpha) &= 2T^2 Q_T \left(\frac{N_T}{T} - \frac{1}{\sqrt{T}} \operatorname{Tr}\left[\mathcal{L}^{-1}_{\frac{R_T}{T}, TQ_T} \left(\frac{M_T}{\sqrt{T}} - 2T^{3/2} Q_T I_d \frac{N_T}{T} \right) \right] \right) \\ \sqrt{T}(\hat{b}_T - b) &= \mathcal{L}^{-1}_{\frac{R_T}{T}, TQ_T} \left(\frac{M_T}{\sqrt{T}} - 2T^{3/2} Q_T I_d \frac{N_T}{T} \right). \end{cases} \tag{30}$$

From (25), we have

$$\frac{N_T}{T} = \frac{1}{2T} \log \left(\frac{\det[X_T]}{\det[x]} \right) - \operatorname{Tr}[b].$$

As for $-b \in S_d^{+,*}$ the Wishart process $(X_t)_{t \geq 0}$ is stationary with invariant limit distribution X_∞ we easily deduce that $\frac{N_T}{T}$ converges in probability to $-\operatorname{Tr}[b]$ when $T \to \infty$. Then, it follows from (19) that

$$(T^{-1} R_T, T^{-1} N_T) \overset{\mathbb{P}_\theta}{\to} (\overline{R}_\infty, -\operatorname{Tr}[b]), \quad as\ T \to \infty. \tag{31}$$

Hence, we only need to study the asymptotic behavior of the couple $(T^{-1/2} M_T, T^2 Q_T)$. According to Theorem 4.1 in Mayerhofer [29], we have for $\lambda \geqslant 0$ and $\Gamma \in S_d$

$$\mathbb{E}_\theta \left[\exp \left(\frac{\lambda}{T} N_T - \frac{\lambda^2}{2T^2} Q_T^{-1} + \frac{1}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] - \frac{2}{T} \int_0^T \operatorname{Tr}[\Gamma^2 X_s] ds - \frac{2\lambda}{\sqrt{T}} \operatorname{Tr}[\Gamma] \right) \right] = 1. \tag{32}$$

Now, let us introduce the quantity

$$\begin{aligned} A_T = \mathbb{E}_\theta \Bigg[&\exp \left(\lambda \frac{N_T}{T} + \lambda \operatorname{Tr}[b] \right) \exp \left(-\frac{\lambda^2}{2T^2} Q_T^{-1} + \frac{1}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] \right) \\ &\times \exp \left(-\frac{2}{T} \int_0^T \operatorname{Tr}[\Gamma^2 X_s] ds + 2 \operatorname{Tr}[\Gamma^2 \overline{R}_\infty] \right) \Bigg]. \end{aligned}$$

Then, by (32) we easily get $A_T = \exp\left(\lambda \operatorname{Tr}[b] + 2 \operatorname{Tr}[\Gamma^2 \overline{R}_\infty] + \frac{2\lambda}{\sqrt{T}} \operatorname{Tr}[\Gamma] \right)$. We now write $A_T = \tilde{A}_T + \mathbb{E}_\theta \left[\exp\left(-\frac{\lambda^2}{2T^2} Q_T^{-1} + \frac{1}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] \right) \right]$ with

$$\begin{aligned} \tilde{A}_T &= \mathbb{E}_\theta \left[(\exp(\xi_T) - 1) \exp \left(-\frac{\lambda^2}{2T^2} Q_T^{-1} + \frac{1}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] \right) \right] \\ \xi_T &= \lambda \frac{N_T}{T} + \lambda \operatorname{Tr}[b] - \frac{2}{T} \int_0^T \operatorname{Tr}[\Gamma^2 X_s] ds + 2 \operatorname{Tr}[\Gamma^2 \overline{R}_\infty]. \end{aligned}$$

Cauchy-Schwarz inequality and $Q_T^{-1} > 0$ give

$$|\tilde{A}_T| \leqslant \mathbb{E}_\theta^{1/2} [\exp(2\xi_T) - 2\exp(\xi_T) + 1] \mathbb{E}_\theta^{1/2} \left[\exp \left(\frac{2}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] \right) \right].$$

On the one hand, Proposition 5.1 with $m = -b \in \mathcal{S}_d^{+,*}$ gives

$$\mathbb{E}_\theta \left[\exp \left(\frac{2}{\sqrt{T}} \operatorname{Tr}[\Gamma M_T] \right) \right] \leqslant \mathbb{E}_\theta \left[\exp \left(\frac{2}{T} \operatorname{Tr}[\Gamma^2 R_T] \right) \right] < \infty.$$

On the other hand, we have for any $r \geqslant 0$,

$$\mathbb{E}_\theta[\exp(r\xi_T)] \leqslant \mathbb{E}_\theta\left[\exp\left(\frac{\lambda r}{T}N_T\right)\right]\exp(2r\operatorname{Tr}[\Gamma^2\overline{R}_\infty]).$$

From (25), we have

$$\mathbb{E}_\theta\left[\exp\left(\frac{\lambda r}{T}N_T\right)\right] = \exp(-\lambda r\operatorname{Tr}[b])\mathbb{E}_\theta\left[\left(\frac{\det[X_T]}{\det[x]}\right)^{\frac{\lambda r}{2T}}\right].$$

The sublinear growth of the coefficients of the Wishart SDE and the convergence to a stationary law gives that $\mathbb{E}_\theta\left[\left(\frac{\det[X_T]}{\det[x]}\right)^{\tilde{\lambda}}\right]$ is uniformly bounded in $T > 0$, $\tilde{\lambda} < 1$ and therefore $\sup_{T>\frac{\lambda r}{2}}\mathbb{E}_\theta\left[\left(\frac{\det[X_T]}{\det[x]}\right)^{\frac{\lambda r}{2T}}\right] < \infty$. This gives the uniform integrability of the family $(\exp(2\xi_T), T > \lambda)$. Then, we deduce from (31) that $\mathbb{E}_\theta[\exp(2\xi_T) - 2\exp(\xi_T) + 1] \underset{T\to+\infty}{\to} 0$ and thus $\tilde{A}_T \underset{T\to+\infty}{\to} 0$.

Hence, we obtain

$$\lim_{T\to\infty}\mathbb{E}_\theta\left[\exp\left(-\frac{\lambda^2}{2T^2}Q_T^{-1} + \frac{1}{\sqrt{T}}\operatorname{Tr}[\Gamma M_T]\right)\right] = \lim_{T\to\infty}A_T = \exp\left(\lambda\operatorname{Tr}[b] + 2\operatorname{Tr}[\Gamma^2\overline{R}_\infty]\right).$$

Therefore, we deduce by Lemma B.4 the following convergence in law

$$\left(\frac{Q_T^{-1}}{T^2}, \frac{M_T}{\sqrt{T}}\right) \Rightarrow \left(\tau_{-\operatorname{Tr}[b]}, \sqrt{\overline{R}_\infty}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top\sqrt{\overline{R}_\infty}\right) \text{ as } T \to \infty,$$

where $\tilde{G}_{i,j}$ $1 \leqslant i,j \leqslant d$ are independent standard normal variables. Together with (31), we obtain that

$$(T^{-1}R_T, T^2Q_T, T^{-1}N_T, T^{-1/2}M_T) \Rightarrow (\overline{R}_\infty, 1/\tau_{-\operatorname{Tr}[b]}, -\operatorname{Tr}[b], \sqrt{\overline{R}_\infty}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top\sqrt{\overline{R}_\infty}), \quad (33)$$

which gives the claim by (30) and Lemma B.4. □

2.2 The global MLE estimator of $\theta = (b, \alpha)$ when $b \in \mathcal{M}_d$

We define the linear operators $\bar{\mathcal{L}}_X, \bar{\mathcal{L}}_{X,a} : \mathcal{M}_d \to \mathcal{M}_d$ by

$$\bar{\mathcal{L}}_X(Y) = \mathcal{L}_X^{-1}(YX + XY^\top)X,\ \bar{\mathcal{L}}_{X,a}(Y) = \bar{\mathcal{L}}_X(Y) - a\operatorname{Tr}[Y]I_d.$$

From (4), we get $Z_T = (\alpha - 1 - d)Q_T^{-1} + 2\operatorname{Tr}[b]T + 2N_T$. This yields with (10) to

$$\begin{cases} \hat{\alpha}_T - \alpha = 2TQ_T\operatorname{Tr}[b - \hat{b}_T] + 2Q_TN_T \\ \int_0^T \bar{\mathcal{L}}_{X_s}(\hat{b}_T - b)ds - T^2Q_T\operatorname{Tr}[\hat{b}_T - b]I_d = \int_0^T \mathcal{L}_{X_s}^{-1}(dM_s)X_s - TQ_TN_TI_d. \end{cases} \quad (34)$$

We now define

$$\hat{\mathcal{L}}_T(Y) = \frac{1}{T}\int_0^T \bar{\mathcal{L}}_{X_s}(Y)ds - TQ_T\operatorname{Tr}[Y]I_d,$$

which is a linear operator on $\mathcal{M}_d$. By using the convexity of the inverse function, there exists $\gamma \in (0,1)$ that depends on $(X_s, s \in [0,T])$ such that $TQ_T = \frac{\gamma}{T}\int_0^T \frac{1}{\mathrm{Tr}[X_s^{-1}]}ds$. We get $\hat{\mathcal{L}}_T(Y) = \frac{1}{T}\int_0^T \bar{\mathcal{L}}_{X_s, \frac{\gamma}{\mathrm{Tr}[X_s^{-1}]}}(Y)ds$. By Lemma B.3, $\hat{\mathcal{L}}_T$ is self adjoint and positive. It is even positive definite since $\mathrm{Tr}[\hat{\mathcal{L}}_T(Y)^\top Y] = 0$ implies by using Lemmas B.3 that $YX_s + X_sY^\top = 0$ a.s. on $[0,T]$ under $\mathbb{P}_\theta$, and therefore the quadratic variation of $\mathrm{Tr}[uYX_s]$ is equal to zero for any $u \in \mathcal{S}_d$. This gives $\int_0^T \mathrm{Tr}[(uY + Y^\top u)X_s(uY + Y^\top u)]ds = 0$ and thus $uY + Y^\top u = 0$ for all $u \in \mathcal{S}_d$, which necessarily implies $Y = 0$. Thus, we rewrite (34) as

$$\begin{cases} \sqrt{T}(\hat{\alpha}_T - \alpha) &= -2TQ_T \,\mathrm{Tr}[\sqrt{T}(\hat{b}_T - b)] + 2TQ_T\frac{N_T}{\sqrt{T}} \\ \sqrt{T}(\hat{b}_T - b) &= \hat{\mathcal{L}}_T^{-1}\left(\frac{1}{\sqrt{T}}\int_0^T \mathcal{L}_{X_s}^{-1}(dM_s)X_s - TQ_T\frac{N_T}{\sqrt{T}}I_d\right). \end{cases} \tag{35}$$

We will assume $-(b+b^\top) \in \mathcal{S}_d^{+,*}$ and know from Lemma C.1 that X_T converges in law under $\mathbb{P}_\theta$ to the stationary law $X_\infty \sim WIS_d(0,\alpha,0,\sqrt{2q_\infty};1/2)$. We define

$$\hat{\mathcal{L}}_\infty(Y) = \mathbb{E}_\theta[\bar{\mathcal{L}}_{X_\infty}(Y)] - \overline{Q}_\infty \,\mathrm{Tr}[Y]I_d. \tag{36}$$

Note that for $Y \in \mathcal{S}_d$, $\hat{\mathcal{L}}_\infty(Y) = Y\overline{R}_\infty - \overline{Q}_\infty \,\mathrm{Tr}[Y]I_d$. From the convexity of the inverse function, $\overline{Q}_\infty = \gamma\mathbb{E}_\theta[1/\,\mathrm{Tr}[X_\infty^{-1}]]$ with $\gamma \in (0,1)$, and thus $\hat{\mathcal{L}}_\infty(Y) = \mathbb{E}_\theta[\bar{\mathcal{L}}_{X_\infty, \frac{\gamma}{\mathrm{Tr}[X_\infty^{-1}]}}(Y)]$ is a self-adjoint positive operator by Lemma B.3. It is even positive definite since $\mathrm{Tr}[\hat{\mathcal{L}}_\infty(Y)^\top Y] = 0$ implies by Lemma B.3 that $YX_\infty + X_\infty Y^\top = 0$ almost surely. Since the law X_∞ has a positive density on $\mathcal{S}_d^{+,*}$, this gives $Yu + uY^\top = 0$ for any $u \in \mathcal{S}_d$ and thus $Y = 0$.

Theorem 2.3. *Assume $-(b+b^\top) \in S_d^{+,*}$ and $\alpha > d+1$. Then, under $\mathbb{P}_\theta$, $\left(\sqrt{T}(\hat{b}_T - b), T(\hat{\alpha}_T - \alpha)\right)$ converges in law when $T \to +\infty$ to the centered Gaussian vector $(\mathbf{G}, H)$ that takes values in $\mathcal{M}_d \times \mathbb{R}$ and has the following Laplace transform: for $c, \lambda \in \mathcal{M}_d \times \mathbb{R}$,*

$$\begin{aligned} \mathbb{E}_\theta[\exp(\mathrm{Tr}[c^\top \mathbf{G}] + \lambda H)] = \exp\Bigg(&\frac{2\overline{Q}_\infty\lambda^2}{1 - \overline{Q}_\infty \,\mathrm{Tr}[\overline{R}_\infty^{-1}]} - \frac{2\overline{Q}_\infty\lambda}{1 - \overline{Q}_\infty \,\mathrm{Tr}[\overline{R}_\infty^{-1}]}\,\mathrm{Tr}[c\overline{R}_\infty^{-1}] \\ &+ \frac{1}{4}\mathbb{E}_\theta[\mathrm{Tr}[\mathcal{L}_{X_\infty}^{-1}(\hat{\mathcal{L}}_\infty^{-1}(c)X_\infty + X_\infty\hat{\mathcal{L}}_\infty^{-1}(c)^\top)(\hat{\mathcal{L}}_\infty^{-1}(c)X_\infty + X_\infty\hat{\mathcal{L}}_\infty^{-1}(c)^\top)]] \\ &+ \frac{\mathrm{Tr}[c\overline{R}_\infty^{-1}]\overline{Q}_\infty}{1 - \overline{Q}_\infty \,\mathrm{Tr}[\overline{R}_\infty^{-1}]}\left(\frac{1}{2}\frac{\mathrm{Tr}[c\overline{R}_\infty^{-1}]}{1 - \overline{Q}_\infty \,\mathrm{Tr}[\overline{R}_\infty^{-1}]} - \mathrm{Tr}[\hat{\mathcal{L}}_\infty^{-1}(c)]\right)\Bigg) \end{aligned} \tag{37}$$

Proof. From the ergodic Birkhoff's theorem, $\hat{\mathcal{L}}_T$ converges almost surely to $\hat{\mathcal{L}}_\infty$, and thus $\hat{\mathcal{L}}_T^{-1}$ converges almost surely to $\hat{\mathcal{L}}_\infty^{-1}$. We define the martingale $\hat{M}_T = \int_0^T \mathcal{L}_{X_s}^{-1}(dM_s)X_s$. We have

$$\begin{aligned} d(\hat{M}_s)_{i,j} &= \mathrm{Tr}[e_{i,j}^\top \mathcal{L}_{X_s}^{-1}(dM_s)X_s] = \frac{1}{2}\,\mathrm{Tr}[\mathcal{L}_{X_s}^{-1}(e_{i,j}X_s + X_se_{i,j}^\top)dM_s] \\ &= \frac{1}{2}\sum_{1\le k,l\le d}(\mathcal{L}_{X_s}^{-1}(e_{i,j}X_s + X_se_{i,j}^\top))_{k,l}(dM_s)_{k,l}. \end{aligned}$$

We get from (27) that

$$\begin{aligned} \langle d(\hat{M}_s)_{i,j}, d(\hat{M}_s)_{i',j'}\rangle &= \frac{1}{2}\,\mathrm{Tr}\left[\mathcal{L}_{X_s}^{-1}(e_{i,j}X_s + X_se_{i,j}^\top)(e_{i',j'}X_s + X_se_{i',j'}^\top)\right]ds, \\ \langle d(\hat{M}_s)_{i,j}, dN_s\rangle &= \mathrm{Tr}[\mathcal{L}_{X_s}^{-1}(e_{i,j}X_s + X_se_{i,j}^\top)]ds = \delta_{i,j}ds. \end{aligned}$$

From the central limit theorem for martingales, $(\frac{\hat{M}_T}{\sqrt{T}}, \frac{N_T}{\sqrt{T}})$ converges in law under $\mathbb{P}_\theta$ towards a centered Gaussian vector $(\hat{\mathbf{G}}, \hat{H})$ taking values in $\mathcal{M}_d \times \mathbb{R}$ such that

$$\begin{aligned}&\mathbb{E}_\theta(\hat{\mathbf{G}}_{i,j}\hat{\mathbf{G}}_{i',j'}) = \frac{1}{2}\mathbb{E}_\theta\left[\operatorname{Tr}\left[\mathcal{L}^{-1}_{X_\infty}(e_{i,j}X_\infty + X_\infty e^\top_{i,j})(e_{i',j'}X_\infty + X_\infty e^\top_{i',j'})\right]\right] \\ &\mathbb{E}_\theta(\hat{\mathbf{G}}_{i,j}\hat{H}) = \delta_{i,j} \text{ and } \mathbb{E}_\theta(\hat{H}^2) = \overline{Q}_\infty^{-1}.\end{aligned} \tag{38}$$

We thus have the following Laplace transform for $c \in \mathcal{M}_d$, $\lambda \in \mathbb{R}$,

$$\mathbb{E}_\theta[\exp(\operatorname{Tr}[c^\top \hat{\mathbf{G}}] + \lambda \hat{H})] = \exp\left(\frac{1}{4}\mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}(cX_\infty + X_\infty c^\top)(cX_\infty + X_\infty c^\top)]] + \lambda \operatorname{Tr}[c] + \frac{1}{2}\lambda^2 \overline{Q}_\infty^{-1}\right). \tag{39}$$

By using Slutsky's Theorem, we get from (35) that $\left(\sqrt{T}(\hat{b}_T - b, \hat{\alpha}_T - \alpha)\right)$ converges in law under $\mathbb{P}_\theta$ when $T \to +\infty$ to the centered Gaussian vector

$$(\mathbf{G}, H) = \left(\hat{\mathcal{L}}_\infty^{-1}(\hat{\mathbf{G}} - \overline{Q}_\infty \hat{H} I_d), 2\overline{Q}_\infty\left(\hat{H} - \operatorname{Tr}[\hat{\mathcal{L}}_\infty^{-1}(\hat{\mathbf{G}} - \overline{Q}_\infty \hat{H} I_d)]\right)\right).$$

Now, we use that $\hat{\mathcal{L}}_\infty^{-1}(I_d) = \frac{1}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]}\overline{R}_\infty^{-1}$ and that $\hat{\mathcal{L}}_\infty$ is self-adjoint to get for $c \in \mathcal{M}_d$,

$$\operatorname{Tr}[c^\top \mathbf{G}] = \operatorname{Tr}[\hat{\mathcal{L}}_\infty^{-1}(c)^\top \hat{\mathbf{G}}] - \frac{\overline{Q}_\infty \hat{H} \operatorname{Tr}[c^\top \overline{R}_\infty^{-1}]}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]},\ H = \frac{2\overline{Q}_\infty}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]}(\hat{H} - \operatorname{Tr}[\overline{R}_\infty^{-1}\hat{\mathbf{G}}])].$$

From (39), we obtain after some calculations (37), using in particular that for $m \in \mathcal{M}_d, s \in \mathcal{S}_d$,

$$\begin{aligned}&\mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}((m+s)X_\infty + X_\infty(m+s)^\top)((m+s)X_\infty + X_\infty(m+s)^\top)]] \\ &= \mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}(mX_\infty + X_\infty m^\top)(mX_\infty + X_\infty m^\top)]] + 2\operatorname{Tr}[s(m\overline{R}_\infty + \overline{R}_\infty m^\top)] + 2\operatorname{Tr}[s^2\overline{R}_\infty]\end{aligned} \tag{40}$$

and taking $m = \hat{\mathcal{L}}_\infty^{-1}(c)$ and $s = -\lambda \frac{2\overline{Q}_\infty}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]}\overline{R}_\infty^{-1}$. □

Remark 2.1. *It is interesting to compare Theorems 2.1 and 2.3 and see that the asymptotic variance of $\sqrt{T}(\hat{\alpha}_T - \alpha)$ is the same in both cases. Instead, for the estimation of the symmetric part of b, we can check that the asymptotic variance is greater when we do not know a priori that b is symmetric. For $c \in \mathcal{S}_d$, we have $c = \frac{1}{2}[\hat{\mathcal{L}}_\infty^{-1}(c)\overline{R}_\infty + \overline{R}_\infty \hat{\mathcal{L}}_\infty^{-1}(c)^\top] - \overline{Q}_\infty \operatorname{Tr}[\hat{\mathcal{L}}_\infty^{-1}(c)]I_d$ and $c = \mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)\overline{R}_\infty + \overline{R}_\infty \mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c) - 2\overline{Q}_\infty \operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)]I_d$. Multiplying by $\overline{R}_\infty^{-1}$, we get*

$$\operatorname{Tr}[\hat{\mathcal{L}}_\infty^{-1}(c)] = \frac{\operatorname{Tr}[c\overline{R}_\infty^{-1}]}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]} = 2\operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)],$$

and then $\hat{\mathcal{L}}_\infty^{-1}(c) = 2\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c) + \Delta$ with $\Delta \overline{R}_\infty + \overline{R}_\infty \Delta^\top = 0$. This gives from (40)

$$\begin{aligned}&\frac{1}{4}\mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}(\hat{\mathcal{L}}_\infty^{-1}(c)X_\infty + X_\infty \hat{\mathcal{L}}_\infty^{-1}(c)^\top)(\hat{\mathcal{L}}_\infty^{-1}(c)X_\infty + X_\infty \hat{\mathcal{L}}_\infty^{-1}(c)^\top)]] \\ &\quad + \frac{\operatorname{Tr}[c\overline{R}_\infty^{-1}]\overline{Q}_\infty}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]}\left(\frac{1}{2}\frac{\operatorname{Tr}[c\overline{R}_\infty^{-1}]}{1 - \overline{Q}_\infty \operatorname{Tr}[\overline{R}_\infty^{-1}]} - \operatorname{Tr}[\hat{\mathcal{L}}_\infty^{-1}(c)]\right) \\ &= 2\operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)^2\overline{R}_\infty] + \frac{1}{4}\mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}(\Delta X_\infty + X_\infty \Delta^\top)(\Delta X_\infty + X_\infty \Delta^\top)]] - 2\overline{Q}_\infty \operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)]^2 \\ &= \operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)c] + \frac{1}{4}\mathbb{E}_\theta[\operatorname{Tr}[\mathcal{L}^{-1}_{X_\infty}(\Delta X_\infty + X_\infty \Delta^\top)(\Delta X_\infty + X_\infty \Delta^\top)]] \geq \operatorname{Tr}[\mathcal{L}^{-1}_{\overline{R}_\infty, \overline{Q}_\infty}(c)c],\end{aligned}$$

since $\mathcal{L}^{-1}_{X_\infty}$ is a self-adjoint positive operator.

2.3 The MLE estimator of b

When $\alpha \in [d-1, d+1)$, we are no longer able to study the convergence of the MLE of α. It is however still possible to get the speed of convergence of the MLE of b.

Theorem 2.4. *Assume that $-b \in S_d^{+,*}$ and $\alpha \geqslant d-1$. For $T > 0$, we consider $\hat{b}_T$ defined by (18). Then, under $\mathbb{P}_\theta$, $\sqrt{T}(\hat{b}_T - b)$ converges in law to a centered Gaussian vector $\mathbf{G}$ on $\mathcal{S}_d$ with the following Laplace transform $\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c\mathbf{G}]\right)\right] = \exp\left(\mathrm{Tr}[c\mathcal{L}^{-1}_{\overline{R}_\infty}(c)]\right)$, $c \in \mathcal{S}_d$.*

Assume that $-(b+b^\top) \in S_d^{+,*}$ and $\alpha \geqslant d-1$. For $T > 0$, we consider $\hat{b}_T$ defined by (16). Then, under $\mathbb{P}_\theta$, $\sqrt{T}(\hat{b}_T - b)$ converges in law to a centered Gaussian vector $\mathbf{G}$ on $\mathcal{M}_d$ with the following Laplace transform: for $c \in \mathcal{M}_d$, $\mathbb{E}_\theta\left[\exp\left(\mathrm{Tr}[c^\top \mathbf{G}]\right)\right]$

$$= \exp\left(\frac{1}{4}\mathbb{E}_\theta[\mathrm{Tr}[\mathcal{L}^{-1}_{X_\infty}(\check{\mathcal{L}}^{-1}_\infty(c)X_\infty + X_\infty\check{\mathcal{L}}^{-1}_\infty(c)^\top)(\check{\mathcal{L}}^{-1}_\infty(c)X_\infty + X_\infty\check{\mathcal{L}}^{-1}_\infty(c)^\top)]]\right),$$

with $\check{\mathcal{L}}_\infty(c) = \mathbb{E}_\theta[\bar{\mathcal{L}}_{X_\infty}(Y)]$.

Proof. We could prove the result for (18) by using the explicit Laplace transform Proposition 5.1. Here, we use the same arguments as before based on the ergodic property. From (18), we have

$$\sqrt{T}(\hat{b}_T - b) = \mathcal{L}^{-1}_{\frac{R_T}{T}}\left(\frac{1}{\sqrt{T}}\left(X_T - x - bR_T - R_T b - \alpha T I_d\right)\right) = \mathcal{L}^{-1}_{\frac{R_T}{T}}\left(\frac{M_T}{\sqrt{T}}\right).$$

As in the proof of Theorem 2.1, $\frac{M_T}{\sqrt{T}}$ converges in law to the centered Gaussian vector $\tilde{\mathbf{G}}$ defined by (28). Slutsky's theorem and (19) give then the convergence of $\sqrt{T}(\hat{b}_T - b)$ to $\mathbf{G} = \mathcal{L}^{-1}_{\overline{R}_\infty}\left(\tilde{\mathbf{G}}\right)$, whose Laplace transform is given by Lemma B.4.

Similarly, we get from (16) that $\sqrt{T}(\hat{b}_T - b) = \check{\mathcal{L}}_T^{-1}(\hat{M}_T/\sqrt{T})$ with $\check{\mathcal{L}}_T(Y) = \frac{1}{T}\int_0^T \bar{\mathcal{L}}_{X_s}(Y)ds$. We easily check that $\mathrm{Tr}[\check{\mathcal{L}}_T(Y)Y] \geq \mathrm{Tr}[\hat{\mathcal{L}}_T(Y)Y]$ and $\mathrm{Tr}[\check{\mathcal{L}}_\infty(Y)Y] \geq \mathrm{Tr}[\hat{\mathcal{L}}_\infty(Y)Y]$ for $Y \in \mathcal{M}_d$. Therefore, $\check{\mathcal{L}}_T$ and $\check{\mathcal{L}}_\infty$ are positive definite and thus invertible. Besides, the ergodic theorem gives that $\check{\mathcal{L}}_T^{-1}$ converges almost surely to $\check{\mathcal{L}}_\infty$. The result follows from (39) and the self-adjoint property of $\check{\mathcal{L}}_\infty$. □

3 Statistical Inference of the Wishart process: some nonergodic cases

This section studies the convergence of the MLE in the case $b = b_0 I_d$ with $b_0 \geqslant 0$. When $b_0 = 0$ and $\alpha \geqslant d+1$, we are able to describe the rate of convergence of the MLE of (b, α) given by (15), when b is known to be symmetric. When $b_0 > 0$ and $\alpha \geqslant d-1$, we can also obtain the rate of convergence of the MLE of b given by (18). Last, when b is known a priori to be diagonal, the MLE of b has a simpler form and we can describe precisely its convergence.

3.1 The global MLE of $\theta = (b, \alpha)$ when $b = 0$

The following result provides the asymptotic behavior of the estimator of the couple when $\alpha > d+1$ and $b = 0$ in (4).

Theorem 3.1. *Assume that $b = 0$ and $\alpha > d+1$. Let $(\hat{b}_T, \hat{\alpha}_T)$ be the MLE defined by (15). Then, $(T(\hat{b}_T - b), \sqrt{\log(T)}(\hat{\alpha}_T - \alpha))$ converges in law under $\mathbb{P}_\theta$ when $T \to +\infty$ to*

$$\left(\mathcal{L}^{-1}_{R^0_1}\left(X^0_1 - \alpha I_d\right), 2\sqrt{\frac{\alpha-(d+1)}{d}}G\right),$$

where $X^0_t = \alpha t I_d + \int_0^t \sqrt{X^0_s}dW_s + dW_s^\top \sqrt{X^0_s}$ is a Wishart process with the same parameters but starting from 0, $R^0_t = \int_0^t X^0_s ds$ and $G \sim \mathcal{N}(0,1)$ is an independent standard Normal variable.

Proof. From (15) and (26), we obtain

$$\begin{cases} \sqrt{\log(T)}(\hat{\alpha}_T - \alpha) &= -2\frac{T\,\mathrm{Tr}[\hat{b}_T]}{\sqrt{\log(T)}}\log(T)Q_T + 2\log(T)Q_T\frac{N_T}{\sqrt{\log(T)}} \\ T\hat{b}_T &= \mathcal{L}^{-1}_{\frac{R_T}{T^2},Q_T}\left(\frac{X_T}{T} - \frac{x}{T} - (Q_T Z_T + 1 + d)I_d\right), \end{cases}$$

and we are interested in studying the convergence in law of $\left(\frac{N_T}{\sqrt{\log(T)}}, \frac{X_T}{T}, \frac{R_T}{T^2}\right)$. By Theorem 4.1 in [29], for $\mu \geqslant 0$ and $T > 1$,

$$\frac{d\overline{\mathbb{P}}}{d\mathbb{P}_\theta} = \exp\left(\frac{\mu N_T}{\sqrt{\log(T)}} - \frac{\mu^2}{2Q_T\log(T)}\right)$$

defines a change of probability and $(X_t)_{t\in[0,T]}$ is a Wishart process with degree $\alpha + \frac{\mu}{\sqrt{\log(T)}}$ under $\overline{\mathbb{P}}$. Let $\lambda_1, \lambda_2 \in \mathcal{S}_d^{+,*}$ and

$$A_T = \mathbb{E}_\theta\left[\exp\left(\frac{\mu N_T}{\sqrt{\log(T)}} - \frac{\mu^2}{2Q_T\log(T)}\right)\exp\left(-\mathrm{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \mathrm{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]\right)\right].$$

By Proposition 5.1, we have

$$\begin{aligned} A_T &= \overline{\mathbb{E}}\left[\exp\left(-\mathrm{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \mathrm{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]\right)\right] \\ &= \frac{1}{\det[V]^{\frac{\alpha+\mu/\sqrt{\log(T)}}{2}}}\exp\left(-\frac{1}{2T}\mathrm{Tr}\left[V'V^{-1}x\right]\right) \underset{T\to+\infty}{\to} \frac{1}{\det[V]^{\frac{\alpha}{2}}}, \end{aligned} \tag{41}$$

where

$$V = (\sqrt{2\lambda_1})^{-1}\sinh(\sqrt{2\lambda_1})2\lambda_2 + \cosh(\sqrt{2\lambda_1}),\ V' = 2\cosh(\sqrt{2\lambda_1})\lambda_2 + \sqrt{2\lambda_1}\sinh(\sqrt{2\lambda_1}). \tag{42}$$

We note that this limit does not depend on μ and is the Laplace transform of (X^0_1, R^0_1) by Proposition 5.1.

We now use that $\frac{1}{Q_T\log(T)} \underset{T\to+\infty}{\to} \frac{d}{\alpha-(d+1)}$ a.s., see Lemma C.2 and we define

$$\tilde{A}_T = \mathbb{E}_\theta\left[\exp(\xi_T)\right] \text{ and } \xi_T = \frac{\mu N_T}{\sqrt{\log(T)}} - \frac{\mu^2 d}{2(\alpha-(d+1))} - \mathrm{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \mathrm{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]$$

that is finite by using equation (73) of Lemma C.2 since $\xi_T \leqslant \frac{\mu N_T}{\sqrt{\log(T)}}$. We have

$$A_T = \tilde{A}_T + \mathbb{E}_\theta \left[\left\{ \exp\left(\frac{\mu^2 d}{2(\alpha-(d+1))} - \frac{\mu^2}{2Q_T \log(T)} \right) - 1 \right\} \exp\left(\xi_T\right) \right].$$

The Cauchy-Schwarz inequality gives

$$(A_T - \tilde{A}_T)^2 \leqslant \mathbb{E}_\theta \left[\left\{ \exp\left(\frac{\mu^2 d}{2(\alpha-(d+1))} - \frac{\mu^2}{2Q_T \log(T)} \right) - 1 \right\}^2 \right] \mathbb{E}_\theta \left[\exp\left(2\xi_T\right)\right].$$

Since $Q_T \log(T)$ is positive for $T > 1$ and converges a.s. to $\frac{\alpha-(d+1)}{d}$, the first expectation goes to 0 while the second one is bounded by using again (73). Therefore, $A_T - \tilde{A}_T \underset{T\to+\infty}{\to} 0$, and we get

$$\mathbb{E}_\theta \left[\exp\left(\frac{\mu N_T}{\sqrt{\log(T)}} - \mathrm{Tr}\left[\frac{\lambda_2}{T} X_T\right] - \mathrm{Tr}\left[\frac{\lambda_1}{T^2} R_T\right] \right) \right] \underset{T\to+\infty}{\to} \frac{\exp\left(\frac{\mu^2 d}{2(\alpha-(d+1))}\right)}{\det[V]^{\frac{\alpha}{2}}}.$$

Thus, $\left(\frac{N_T}{\sqrt{\log(T)}}, \frac{X_T}{T}, \frac{R_T}{T^2}\right)$ converges in law to $(\sqrt{\frac{d}{\alpha-(d+1)}}G, X_1^0, R_1^0)$, where $G \sim \mathcal{N}(0,1)$ is independent of X^0. From (25), we have

$$Q_T Z_T + 1 + d = 2\frac{1}{\sqrt{\log(T)}} \log(T) Q_T \frac{N_T}{\sqrt{\log(T)}} + \alpha,$$

and therefore $Q_T Z_T + 1 + d$ converges in probability to α. Slutsky's theorem gives then the following convergence in law: as $T \to +\infty$,

$$\left(\frac{N_T}{\sqrt{\log(T)}}, \frac{X_T}{T}, \frac{R_T}{T^2}, Q_T Z_T + 1 + d, Q_T \log(T) \right) \Rightarrow \left(\sqrt{\frac{d}{\alpha-(d+1)}} G, X_1^0, R_1^0, \alpha, \frac{\alpha-(d+1)}{d} \right). \tag{43}$$

This gives the claimed convergence for $(\hat{\alpha}_T, \hat{b}_T)$ due to the continuity property given in Lemma B.1. □

Theorem 3.2. *Assume that $b = 0$ and $\alpha = d+1$. Let $(\hat{b}_T, \hat{\alpha}_T)$ be the MLE defined by (15). Then, $(T(\hat{b}_T - b), \log(T)(\hat{\alpha}_T - \alpha))$ converges in law under $\mathbb{P}_\theta$ when $T \to +\infty$ to*

$$\left(\mathcal{L}_{R_1^0}^{-1}\left(X_1^0 - \alpha I_d\right), \frac{4}{d\tau_1} \right),$$

where $X_t^0 = \alpha t I_d + \int_0^t \sqrt{X_s^0} dW_s + dW_s^\top \sqrt{X_s^0}$ is a Wishart process with the same parameters but starting from 0, $R_t^0 = \int_0^t X_s^0 ds$ and $\tau_1 = \inf\{t \geqslant 0, B_t = 1\}$ where B is a standard Brownian motion independent from W.

Proof. The proof follows the same line as the one of Theorem 3.1, but we now write

$$\log(T)(\hat{\alpha}_T - \alpha) = -2\frac{T\,\mathrm{Tr}[\hat{b}_T]}{\log(T)} \log(T)^2 Q_T + 2\log(T)^2 Q_T \frac{N_T}{\log(T)},$$

while we still have $T\hat{b}_T = \mathcal{L}^{-1}_{\frac{R_T}{T^2},Q_T}\left(\frac{X_T}{T} - \frac{x}{T} - (Q_T Z_T + 1 + d)I_d\right)$. By Theorem 4.1 in [29], for $\mu \geqslant 0$ and $T > 1$, $\frac{d\overline{\mathbb{P}}}{d\mathbb{P}} = \exp\left(\frac{\mu N_T}{\log(T)} - \frac{\mu^2}{2Q_T\log(T)^2}\right)$ defines a change of probability, and we define for $\lambda_1, \lambda_2 \in \mathcal{S}_d^{+,*}$,

$$A_T = \mathbb{E}_\theta\left[\exp\left(\frac{\mu N_T}{\log(T)} - \frac{\mu^2}{2Q_T\log(T)^2}\right)\exp\left(-\operatorname{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \operatorname{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]\right)\right].$$

By Proposition 5.1, we have

$$A_T = \frac{1}{\det[V]^{\frac{\alpha+\mu/\log(T)}{2}}}\exp\left(-\frac{1}{2T}\operatorname{Tr}\left[V'V^{-1}x\right]\right) \underset{T\to+\infty}{\to} \frac{1}{\det[V]^{\frac{\alpha}{2}}},$$

where V and V' are defined by (42).

We now use that $\frac{N_T}{\log(T)} \to \frac{d}{2}$ in probability, see Lemma C.2, and define

$$\tilde{A}_T = \mathbb{E}_\theta\left[\exp(\xi_T)\right],\ \xi_T = \frac{\mu d}{2} - \frac{\mu^2}{2Q_T\log(T)^2} - \operatorname{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \operatorname{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]$$

and have

$$A_T = \tilde{A}_T + \mathbb{E}_\theta\left[\left\{\exp\left(\frac{\mu N_T}{\log(T)} - \frac{\mu d}{2}\right) - 1\right\}\exp\left(\xi_T\right)\right].$$

We note that $\exp(\xi_T) \leqslant \exp\left(\frac{\mu d}{2}\right)$. By using Lemma C.2 and the uniform integrability (74), we get that $A_T - \tilde{A}_T \underset{T\to+\infty}{\to} 0$ and therefore

$$\mathbb{E}_\theta\left[\exp\left(-\frac{\mu^2}{2Q_T\log(T)^2} - \operatorname{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \operatorname{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]\right)\right] \underset{T\to+\infty}{\to} \frac{\exp(-\mu d/2)}{\det[V]^{\frac{\alpha}{2}}}.$$

Therefore, $\left(\frac{X_T}{T}, \frac{R_T}{T^2}, Q_T\log(T)^2\right)$ converges in law to $\left(X_1^0, R_1^0, \left(\frac{2}{d}\right)^2\frac{1}{\tau_1}\right)$, where τ_1 is independent of X^0. We observe that $Q_T Z_T = \frac{1}{\log(T)}Q_T\log(T)^2\frac{Z_T}{\log(T)}$. Lemma C.2 and Slutsky's theorem gives

$$\left(\frac{N_T}{\log(T)}, \frac{X_T}{T}, \frac{R_T}{T^2}, Q_T Z_T + 1 + d, Q_T\log(T)^2\right) \Rightarrow \left(\frac{d}{2}, X_1^0, R_1^0, d+1, \left(\frac{2}{d}\right)^2\frac{1}{\tau_1}\right), \tag{44}$$

which gives the claim by using the formulas for $T\hat{b}_T$ and $\log(T)(\hat{\alpha}_T - \alpha)$. □

3.2 The MLE of b

Until the end of this section we consider that $\alpha \geqslant d-1$ is known and study the speed of convergence of the estimator of b defined by (18).

3.2.1 Case $b = 0$.

Theorem 3.3. *Assume that $b = 0$ and $\alpha \geqslant d-1$. For $T > 0$, let $\hat{b}_T$ be defined by (18). When $T \to +\infty$, $T(\hat{b}_T - b)$ converges in law under $\mathbb{P}_\theta$ to $\mathcal{L}^{-1}_{R_1^0}\left(X_1^0 - \alpha I_d\right)$, where $(X_t^0)_{t\geqslant 0}$ is the solution to $X_t^0 = \alpha t I_d + \int_0^t \sqrt{X_s^0}dW_s + dW_s^\top\sqrt{X_s^0}$ and $R_t^0 = \int_0^t X_s^0 ds$.*

Proof. From (18), we have $T\hat{b}_T = \mathcal{L}^{-1}_{\frac{R_T}{T^2}}\left(\frac{X_T}{T} - \frac{x}{T} - \alpha I_d\right)$. Let V and V' be defined by (42). Similarly to (41), we have by Proposition 5.1 for $\lambda_1, \lambda_2 \in \mathcal{S}_d^{+,*}$

$$\mathbb{E}_\theta\left[\exp\left(-\operatorname{Tr}\left[\frac{\lambda_2}{T}X_T\right] - \operatorname{Tr}\left[\frac{\lambda_1}{T^2}R_T\right]\right)\right] = \frac{1}{\det[V]^{\frac{\alpha}{2}}}\exp\left(-\frac{1}{2T}\operatorname{Tr}\left[V'V^{-1}x\right]\right) \underset{T\to+\infty}{\to} \frac{1}{\det[V]^{\frac{\alpha}{2}}}.$$

This gives the convergence in law of $\left(\frac{X_T}{T}, \frac{R_T}{T^2}\right)$ to (X_1^0, R_1^0) and then the claimed result. □

3.2.2 Case $b = b_0 I_d, b_0 > 0$.

In this case $b = b_0 I_d$ with $b_0 > 0$. In order to identify the speed of convergence and the limit law, we use the Laplace transform approach. We have the following result,

Theorem 3.4. *Assume that $b = b_0 I_d$, $b_0 > 0$, and $\alpha \geqslant d-1$. For $T > 0$ let $\hat{b}_T$ defined by (18). When $T \to +\infty$, $\exp(b_0 T)(\hat{b}_T - b)$ converges in law under $\mathbb{P}_\theta$ to $\mathcal{L}_X^{-1}\left(\sqrt{X}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}\sqrt{X}\right)$ where $X \sim WIS_d\left(\frac{x}{2b_0}, \alpha, 0, I_d; \frac{1}{4b_0^2}\right)$ and $\tilde{\mathbf{G}}$ is an independent d-square matrix whose elements are independent standard Normal variables.*

The proof of this results relies on the explicit calculation of the Laplace transform of (X_T, R_T) and is postponed to Subsection 5.2.

Obviously, the case $b = b_0 I_d$ is very particular. One would like to consider more general nonergodic cases or ideally to be able to state a general convergence results of $\hat{b}_T$ towards b for any $b \in \mathcal{S}_d$. Despite our efforts, we have not been able to get such a result. The reason why we can handle the ergodic case and the nonergodic case $b = b_0 I_d$ with $b_0 \geqslant 0$ is that the convergence of all the matrix terms occurs at the same speed, namely $1/\sqrt{T}$ for the ergodic case, $1/T$ for $b = 0$ and $e^{-b_0 T}$ when $b_0 > 0$. In the other cases, there is no such a simple scalar rescaling. Heuristically, there may be different speeds of convergence that are difficult to disentangle because of the different matrix products. To get an idea of this, we present now the case of the estimation of b when b is known to be a diagonal matrix. In this case, we obtain different speed of convergence for each diagonal terms.

3.2.3 The MLE of b when b is known a priori to be diagonal.

We assume that $\alpha \geqslant d-1$ is known and that b is a diagonal matrix, i.e. $b = \operatorname{diag}(b_1, \cdots, b_d)$. We want to estimate the diagonal elements by maximizing the likelihood. We denote $\theta_0 = (0, \alpha)$. As in (17), we have

$$L_T^{\theta,\theta_0} = \exp\Big(\frac{\operatorname{Tr}[bX_T] - \operatorname{Tr}[bx]}{2} - \frac{1}{2}\int_0^T \operatorname{Tr}[b^2 X_s]ds - \frac{\alpha T}{2}\operatorname{Tr}[b]\Big).$$

By differentiating this with respect to b_i, $1 \leqslant i \leqslant d$, we get

$$\frac{\partial_{b_i} L_T^{\theta,\theta_0}}{L_T^{\theta,\theta_0}} = \frac{1}{2}\Big((X_T)_{i,i} - x_{i,i} - \alpha t I_d - 2b_i\int_0^T (X_s)_{i,i}ds\Big),$$

and therefore the MLE of b is given by

$$(\hat{b}_T)_i = \frac{(X_T)_{i,i} - x_{i,i} - \alpha T}{2(R_T)_{i,i}}. \tag{45}$$

We therefore obtain

$$(\hat{b}_T)_i - b_i = \frac{(X_T)_{i,i} - x_{i,i} - \alpha T - 2b_i(R_T)_{i,i}}{2(R_T)_{i,i}}. \tag{46}$$

Let us observe that this estimator is precisely the one obtained by Ben Alaya and Kebaier [5] for the CIR process. This is not very surprising since we know from (4), (2) and b diagonal that there exists independent Brownian motions β^i, $1 \leqslant i \leqslant d$ such that

$$d(X_t)_{i,i} = (\alpha + 2b_i(X_t)_{i,i})dt + 2\sqrt{(X_t)_{i,i}}d\beta_t^i.$$

Thus, the diagonal elements follow independent CIR processes, and the observation of the non diagonal elements does not improve the ML estimation. We can obtain the asymptotic convergence by applying Theorem 1 in [4], up to a small correction in the nonergodic case which is given by our Theorem 3.4 in dimension $d = 1$. This yields to the following proposition.

Proposition 3.1. *Let $\alpha \geqslant d-1$ and b a diagonal matrix. Let $\epsilon_t = \operatorname{diag}(\epsilon_t^1, \cdots, \epsilon_t^d)$ be a diagonal matrix with*

$$\epsilon_t^i = \begin{cases} t^{-\frac{1}{2}} & \text{if} \quad b_i < 0 \\ t^{-1} & \text{if} \quad b_i = 0 \\ \exp(-b_i t) & \text{if} \quad b_i > 0 \end{cases}$$

Then, under $\mathbb{P}_\theta$, $\epsilon_T^{-1} \operatorname{diag}((\hat{b}_T)_1 - b_1, \cdots, (\hat{b}_T)_d - b_d)$ converges in law to a diagonal matrix $\mathbf{D}$ made with independent elements. Each diagonal element $\mathbf{D}_i$ is distributed as follows:

$$\forall i \in \{1, \cdots, d\}, \quad \mathbf{D}_i \underset{law}{=} \begin{cases} \sqrt{\frac{-2b_i}{\alpha}}\mathbf{G} & \text{if } b_i < 0 \\ \frac{X_1^0 - \alpha}{2R_1^0} & \text{if } b_i = 0 \\ \frac{\mathbf{G}}{\sqrt{X^{x_{i,i}/(2b_i)}_{1/(4b_i^2)}}}, & \text{if } b_i > 0 \end{cases}$$

where $X_t^x = x + \alpha t + 2\int_0^t \sqrt{X_s^x}dW_s$, $R_t^0 = \int_0^t X_s^0 ds$, and $\mathbf{G} \sim \mathcal{N}(0,1)$ is independent of X.

4 Optimality of the MLE

In parametric estimation theory, a fundamental role is played by the local asymptotic normality (**LAN**) property since the work of Le Cam [24]. This general concept developed by Le Cam is extended later by Le Cam and Yang [25] and Jeganathan [22] to local asymptotic mixed normality (**LAMN**) and local asymptotic quadraticity (**LAQ**) properties. These notions are mainly dedicated to study the asymptotic efficiency of estimators of a given parametric model. The aim of this section is to check the validity of either **LAN**, **LAMN** or **LAQ** properties for the global model in order to get the asymptotic efficiency of our maximum likelihood estimators studied in the previous section. Here we prove these properties only for the global model $\theta = (b, \alpha)$ when b is known to be symmetric. The same technique applies for all the other cases considered in this paper where we have been able to obtain the corresponding local asymptotic property.

Let us consider the Wishart process $(X_t)_{t\geqslant 0} \in \mathcal{S}_d^+$ with parameters $\theta := (b,\alpha)$, with $\alpha \geq d+1$ and $b \in \mathcal{S}_d$.

$$\begin{cases} dX_t = [\alpha I + bX_t + X_t b]\,dt + \sqrt{X_t}dW_t + dW_t^\top \sqrt{X_t}, t > 0 \\ X_0 \in \mathcal{S}_d^+. \end{cases} \tag{47}$$

We recall that $\mathbb{P}_\theta$ denotes the distributions induced by the solutions of (47) on canonical space $C(\mathbb{R}_+, \mathcal{S}_d^+)$ with the natural filtration $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$ and $\mathbb{P}_{\theta,t} = \mathbb{P}_{\theta|\mathcal{F}_t^X}$ denotes the restriction of $\mathbb{P}_\theta$ on the filtration $\mathcal{F}_t$.

For $\tilde{\alpha} \geq d+1$ and $\tilde{b} \in \mathcal{S}_d^+$, we set $\tilde{\theta} = (\tilde{b}, \tilde{\alpha})$,

$$H_t = \frac{\tilde{\alpha} - \alpha}{2}(\sqrt{X_t})^{-1} + (\tilde{b} - b)\sqrt{X_t},$$

and we introduce the log-likelihood function

$$\ell_T^\theta(\tilde{\theta}) \quad = \quad \log\left(\frac{d\mathbb{P}_{\tilde{\theta},T}}{d\mathbb{P}_{\theta,T}}\right) = \int_0^T \mathrm{Tr}[H_s dW_s] - \frac{1}{2}\int_0^T \mathrm{Tr}[H_s H_s^\top]ds \tag{48}$$

The process $(\tilde{W}_t = W_t - \int_0^t H_s^\top ds, t \leq T)$ is a $d \times d$-Brownian motion under $\mathbb{P}_{\tilde{\theta},T}$. In the sequel, let us introduce the quantity $\delta_T := (\delta_{1,T}, \delta_{2,T}) \in \mathbb{R}^2$ where for $i \in \{1,2\}$ the localizing rates satisfy $|\delta_{i,T}| \to 0$ when $T \to \infty$. For all $u := (u_1, u_2) \in \mathbb{R} \times \mathcal{S}_d$, we define $\delta_T \cdot u := (\delta_{1,T}u_1, \delta_{2,T}u_2) \in \mathbb{R} \times \mathcal{S}_d$. Now, we rewrite (48) with $\tilde{\theta} = \theta + \delta_T \cdot u$

$$\begin{aligned} \ell_T^\theta(\theta + \delta_T \cdot u) \quad &= \quad \int_0^T \mathrm{Tr}\left[\frac{\delta_{1,T}u_1(\sqrt{X_s})^{-1}}{2}dW_s + \delta_{2,T}u_2\sqrt{X_s}dW_s\right] \\ &\quad -\frac{1}{2}\int_0^T \mathrm{Tr}\left[\delta_{2,T}u_2 X_s \delta_{2,T}u_2\right] ds - \frac{T}{2}\delta_{1,T}u_1 \,\mathrm{Tr}\left[\delta_{2,T}u_2\right] \\ &\quad -\frac{1}{8}\int_0^T (\delta_{1,T}u_1)^2\, \mathrm{Tr}\left[(X_s)^{-1}\right] ds. \end{aligned}$$

Hence, by using the definitions (13), (23) and (24) of the martingales processes $(N_t)_{t\geq 0}$ and $(M_t)_{t\geq 0}$ and the processes $(R_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$, it is easy to check that

$$\begin{aligned} \ell_T^\theta(\theta + \delta_T \cdot u) \quad &= \quad \frac{1}{2}\Big(\delta_{1,T}u_1 N_T + \mathrm{Tr}\big[\delta_{2,T}u_2 M_T\big]\Big) - \frac{1}{2}\mathrm{Tr}\left[\delta_{2,T}u_2 R_T \delta_{2,T}u_2\right] \\ &\quad -\frac{T}{2}\delta_{1,T}u_1\,\mathrm{Tr}\left[\delta_{2,T}u_2\right] - \frac{1}{8}(\delta_{1,T}u_1)^2 (Q_T)^{-1} \\ &= \quad \Lambda_T(u) - \frac{1}{2}\Gamma_T(u), \end{aligned} \tag{49}$$

where $\Lambda_T(u) = \frac{1}{2}\Big(\delta_{1,T}u_1 N_T + \mathrm{Tr}\big[\delta_{2,T}u_2 M_T\big]\Big)$ is a linear random function with respect to $u \in \mathbb{R} \times \mathcal{S}_d^+$ with quadratic variation

$$\Gamma_T(u) = \delta_{2,T}^2\,\mathrm{Tr}[u_2^2 R_T] + T\delta_{1,T}\delta_{2,T}u_1\,\mathrm{Tr}\left[u_2\right] + \frac{1}{4}\delta_{1,T}^2 u_1^2 Q_T^{-1}.$$

4.1 Case $-b \in S_d^+$ and $\alpha \geqslant d+1$

We first consider $\alpha > d+1$. In this ergodic case, we set $\delta_{i,T} = T^{-1/2}$ for $i \in \{1,2\}$, and we get from (19) and (21)

$$\Gamma_T(u) \xrightarrow{a.s.} \overline{\Gamma}_\infty(u) := \operatorname{Tr}\left[u_2^2 \overline{R}_\infty\right] + u_1 \operatorname{Tr}\left[u_2\right] + \frac{1}{4} u_1^2 \operatorname{Tr}\left[(\overline{Q}_\infty)^{-1}\right], \quad \text{as } T \to +\infty. \tag{50}$$

This yields the validity of the so called Raykov type condition. Hence, according to Theorem 1 in [28], relations (49) and (50) ensure the validity of the local asymptotic normality (**LAN**) property, that is under $\mathbb{P}_\theta$ we have

$$(\Lambda_T(u), \Gamma_T(u)) \Rightarrow \left(\overline{\Gamma}_\infty^{1/2}(u) Z, \overline{\Gamma}_\infty(u)\right), \text{ as } T \to \infty, \tag{51}$$

with Z a standard normal real random variable. It is worth noting that the above convergence can also be obtained using the proof of Theorem 2.1. In fact, we have already proven that under $\mathbb{P}_\theta$

$$\left(\frac{N_T}{\sqrt{T}}, \frac{M_T}{\sqrt{T}}\right) \Rightarrow (\tilde{\mathbf{G}}, \tilde{H}) \tag{52}$$

where $(\tilde{G}, \tilde{H})$ is a centered Gaussian vector taking values in $\mathbb{R} \times S_d^+$ such that

$$\begin{aligned}
\mathbb{E}_\theta(\tilde{\mathbf{G}}_{i,j} \tilde{\mathbf{G}}_{k,l}) &= \left[\delta_{jl} (\overline{R}_\infty)_{i,k} + \delta_{jk} (\overline{R}_\infty)_{i,l} + \delta_{il} (\overline{R}_\infty)_{j,k} + \delta_{ik} (\overline{R}_\infty)_{j,l}\right], \\
\mathbb{E}_\theta(\tilde{\mathbf{G}}_{i,j} \tilde{H}) &= 2\delta_{i,j} \text{ and } \mathbb{E}_\theta(\tilde{H}^2) = \overline{Q}_\infty^{-1}.
\end{aligned}$$

Therefore, **LAN** property (51) follows from relations (50) and (52).

We now consider the case $\alpha = d+1$ and set $\delta_{1,T} = T^{-1}$ and $\delta_{2,T} = T^{-1/2}$. By using (33), we get that under $\mathbb{P}_\theta$,

$$(\Lambda_T(u), \Gamma_T(u)) \Rightarrow \left(-\frac{u_1}{2} \operatorname{Tr}[b] + \operatorname{Tr}\left[u_2 \sqrt{\overline{R}_\infty} \tilde{\mathbf{G}}\right], \operatorname{Tr}[u_2^2 \overline{R}_\infty] + \frac{1}{4} u_1^2 \tau_{-\operatorname{Tr}[b]}\right), \text{ as } T \to \infty,$$

where $\tau_{-\operatorname{Tr}[b]}$ is defined as in Theorem 2.2 and $\tilde{G}$ is an independent matrix, whose elements $\tilde{G}_{i,j}$, $1 \leqslant i, j \leqslant d$, are independent standard normal variables. Hence, according to Le Cam and Yang [25] and Jeganathan [22] this last convergence yields the **LAQ** property for this ergodic case.

4.2 Case $b = 0$ and $\alpha \geqslant d+1$

We first assume $\alpha > d+1$. From (47) with $b = 0$ and (23), we have $M_T = X_T - x - \alpha I_d T$. From (43), it follows that as $T \to \infty$

$$\left(\frac{N_T}{\sqrt{\log(T)}}, \frac{M_T}{T}, \frac{R_T}{T^2}, \frac{Q_T^{-1}}{\log(T)}\right) \Rightarrow \left(\sqrt{\frac{d}{\alpha - (d+1)}} G, X_1^0 - \alpha I_d, R_1^0, \frac{d}{\alpha - (d+1)}\right),$$

where X_1^0 and R_1^0 are defined as in Theorem 3.1. Thus, in the same way as in the previous case if we set $\delta_{1,T} = \frac{1}{\sqrt{\log(T)}}$ and $\delta_{2,T} = T^{-1}$, then $(\Lambda_T(u), \Gamma_T(u))$ converges in law under $\mathbb{P}_\theta$ to

$$\left(\frac{1}{2} \sqrt{\frac{d}{\alpha - (d+1)}} u_1 G + \frac{1}{2} \operatorname{Tr}\left[u_2 (X_1^0 - \alpha I_d)\right], \operatorname{Tr}\left[u_2^2 \overline{R}_1^0\right] + \frac{u_1^2 d}{4(\alpha - (d+1))}\right), \text{ as } T \to \infty.$$

This ensures the validity of the **LAQ** property in this non-ergodic case.

When $\alpha = d+1$, we use the notation of Theorem 3.2 and get from (44)

$$\left(\frac{N_T}{\log(T)}, \frac{M_T}{T}, \frac{R_T}{T^2}, \frac{Q_T^{-1}}{\log(T)^2}\right) \Rightarrow \left(\frac{d}{2}, X_1^0 - \alpha I_d, R_1^0, \left(\frac{d}{2}\right)^2 \tau_1\right).$$

With $\delta_{1,T} = \frac{1}{\log(T)}$ and $\delta_{2,T} = T^{-1}$, we get that $(\Lambda_T(u), \Gamma_T(u))$ converges in law under $\mathbb{P}_\theta$ to

$$\left(\frac{d}{4}u_1 + \frac{1}{2}\operatorname{Tr}[u_2(X_1^0 - \alpha I_d)], \operatorname{Tr}\left[u_2^2 \overline{R}_1^0\right] + \frac{d^2 u_1^2}{8}\tau_1\right), \quad \text{as } T \to \infty.$$

This gives again the **LAQ** property.

5 The Laplace transform and its use to study the MLE

5.1 The Laplace transform of (X_T, R_T)

We present our main result on the joint Laplace transform of (X_T, R_T), that can be of independent interest. This Laplace transform is given by Bru [7], eq. (4.7) when $b = 0$ and has been recently studied and obtained explicitly by Gnoatto and Grasselli [17]. Here, we present another proof that enables us to get the Laplace transform for any $\alpha \geqslant d-1$, as well as a more precise result concerning its set of convergence, see Remarks 5.1 and 5.2 below for a further discussion.

Proposition 5.1. *Let $\alpha \geqslant d-1$, $x \in \mathcal{S}_d^+$, $b \in \mathcal{S}_d$ and $X \sim WIS_d(x, \alpha, b, I_d)$. Let $v, w \in \mathcal{S}_d$ be such that*

$$\exists m \in \mathcal{S}_d,\ \frac{v}{2} - mb - bm - 2m^2 \in \mathcal{S}_d^+ \text{ and } \frac{w}{2} + m \in \mathcal{S}_d^+. \tag{53}$$

Then, we have for $t \geqslant 0$

$$\begin{aligned}&\mathbb{E}\Big[\exp\Big(-\frac{1}{2}\operatorname{Tr}[wX_t] - \frac{1}{2}\operatorname{Tr}[vR_t]\Big)\Big] \\ &= \frac{\exp\big(-\frac{\alpha}{2}\operatorname{Tr}[b]t\big)}{\det[V_{v,w}(t)]^{\frac{\alpha}{2}}} \exp\Big(-\frac{1}{2}\operatorname{Tr}\left[(V'_{v,w}(t)V_{v,w}(t)^{-1} + b)x\right]\Big),\end{aligned} \tag{54}$$

with

$$V_{v,w}(t) = \left(\sum_{k=0}^{\infty} t^{2k+1}\frac{\tilde{v}^k}{(2k+1)!}\right)\tilde{w} + \sum_{k=0}^{\infty} t^{2k}\frac{\tilde{v}^k}{(2k)!}, \quad \tilde{v} = v + b^2, \quad \text{and} \quad \tilde{w} = w - b.$$

If besides $\tilde{v} = v + b^2 \in \mathcal{S}_d^{+,*}$, we have $V_{v,w}(t) = (\sqrt{\tilde{v}})^{-1}\sinh(\sqrt{\tilde{v}}t)\tilde{w} + \cosh(\sqrt{\tilde{v}}t)$ and then $V'_{v,w}(t) = \cosh(\sqrt{\tilde{v}}t)\tilde{w} + \sinh(\sqrt{\tilde{v}}t)\sqrt{\tilde{v}}$.

Before proving this result, we recall the following fact:

$$\forall x, y \in \mathcal{S}_d^+,\ \operatorname{Tr}[xy] \geqslant 0, \tag{55}$$

which is clear once we have observed that $\operatorname{Tr}[xy] = \operatorname{Tr}[\sqrt{x}y\sqrt{x}]$ and $\sqrt{x}y\sqrt{x} \in \mathcal{S}_d^+$. We also recall a result on matrix Riccati equations, see Dieci and Eirola [11] Proposition 1.1.

Lemma 5.1. *Let $\tilde{b} \in \mathcal{S}_d$ and $\tilde{\delta} \in \mathcal{S}_d^+$. Let ξ denote the solution of the following matrix Riccati differential equation*

$$\xi' + 2\xi^2 = \tilde{b}\xi + \xi\tilde{b} + \tilde{\delta}, \ \xi(0) \in \mathcal{S}_d. \tag{56}$$

If $\xi(0) \in \mathcal{S}_d^+$, the solution ξ is well-defined for any $t \geqslant 0$ and satisfies $\xi(t) \in \mathcal{S}_d^+$.

Proof of Proposition 5.1. Let $T > 0$ be given. We first assume $w, v \in \mathcal{S}_d^+$, which ensures that $\mathbb{E}\left[\exp\left(-\frac{1}{2}\operatorname{Tr}[wX_T + vR_T]\right)\right] < \infty$. We consider the martingale

$$M_t = \mathbb{E}\left[\exp\left(-\frac{1}{2}\operatorname{Tr}[wX_T + vR_T]\right)\middle|\mathcal{F}_t\right], \ t \in [0,T].$$

Due to the affine structure, we are looking for smooth functions $\beta : \mathbb{R}_+ \to \mathbb{R}$, $\gamma, \delta : \mathbb{R}_+ \to \mathcal{S}_d$ such that

$$M_t = \exp\left(\beta(T-t) + \operatorname{Tr}[\gamma(T-t)X_t] + \operatorname{Tr}[\delta(T-t)R_t]\right).$$

We necessarily have $\beta(0) = 0$, $\gamma(0) = -w/2$ and $\delta(0) = -v/2$. Itô's formula gives

$$\begin{aligned}\frac{dM_t}{M_t} =& \Big\{ -\beta'(T-t) - \operatorname{Tr}[\gamma'(T-t)X_t] - \operatorname{Tr}[\delta'(T-t)R_t] + \operatorname{Tr}[\gamma(T-t)(\alpha I_d + bX_t + X_t b)] \\ &+ \operatorname{Tr}[\delta(T-t)X_t] + 2\operatorname{Tr}[\gamma(T-t)^2X_t]\Big\}dt + \operatorname{Tr}[\gamma(T-t)(\sqrt{X_t}dW_t + dW_t^\top\sqrt{X_t})].\end{aligned}$$

Since M is a martingale, the drift term should vanish almost surely. The drift term being a (deterministic) affine function of (X_t, R_t), we obtain the following system of differential equations:

$$\delta' = 0, \tag{57}$$

$$-\gamma' + \gamma b + b\gamma + 2\gamma^2 + \delta = 0, \tag{58}$$

$$-\beta' + \alpha\operatorname{Tr}[\gamma] = 0. \tag{59}$$

The first equation gives $\delta(t) = -v/2$. The second equation is a matrix Riccati differential equation. We now consider $\xi = m - \gamma$ with m satisfying (53). It solves (56) with $\tilde{b} = b + 2m$, $\tilde{\delta} = -bm - mb - 2m^2 + v/2$ and $\xi(0) = m + w/2$. We know then by Lemma 5.1 that ξ is well defined for any $t \geqslant 0$ and stays in $\mathcal{S}_d^+$. In particular, γ is well defined for any $t \geqslant 0$.

We set $\tilde{\gamma} = \gamma + \frac{1}{2}b$. We have $\gamma^2 = \tilde{\gamma}^2 - \frac{1}{2}(b\tilde{\gamma} + \tilde{\gamma}b) + \frac{1}{4}b^2$ and thus $\tilde{\gamma}$ solves the following matrix Riccati differential equation:

$$\tilde{\gamma}' = 2\tilde{\gamma}^2 - \frac{1}{2}\tilde{v}, \ \tilde{\gamma}(0) = -\frac{1}{2}\tilde{w}, \text{ with } \tilde{v} = v + b^2 \text{ and } \tilde{w} = w - b.$$

We set $M(t) = \begin{bmatrix} M_1(t) & M_2(t) \\ M_3(t) & M_4(t) \end{bmatrix} = \exp\left(t\begin{bmatrix} 0 & -\tilde{v}/2 \\ -2I_d & 0 \end{bmatrix}\right) \in \mathcal{M}_{2d}$ and get by Levin [26] that

$$\tilde{\gamma}(t) = \left[M_2(t) - \frac{1}{2}M_1(t)\tilde{w}\right]\left[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}\right]^{-1}.$$

We check that the matrix $M_4(t) - \frac{1}{2}M_3(t)\tilde{w}$ is indeed invertible. In fact, let

$$\tau = \inf\left\{t \geqslant 0, \det\left[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}\right] = 0\right\}.$$

We have $\tau > 0$ and for $t \in [0,\tau)$, $\frac{d}{dt}[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}] = -2\left[M_2(t) - \frac{1}{2}M_1(t)\tilde{w}\right]$ and thus

$$\frac{d}{dt}\det\left[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}\right] = -2\det\left[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}\right]\mathrm{Tr}[\tilde{\gamma}(t)].$$

This gives $\det\left[M_4(t) - \frac{1}{2}M_3(t)\tilde{w}\right] = \exp(-2\int_0^t \mathrm{Tr}[\tilde{\gamma}(s)]ds) > 0$, and we necessary get $\tau = +\infty$ since γ and thus $\tilde{\gamma}$ is well defined for $t \geqslant 0$.

Since

$$\begin{bmatrix} 0 & -\tilde{v}/2 \\ -2I_d & 0 \end{bmatrix}^{2k} = \begin{bmatrix} \tilde{v}^k & 0 \\ 0 & \tilde{v}^k \end{bmatrix} \text{ and } \begin{bmatrix} 0 & -\tilde{v}/2 \\ -2I_d & 0 \end{bmatrix}^{2k+1} = \begin{bmatrix} 0 & -\tilde{v}^{k+1}/2 \\ -2\tilde{v}^k & 0 \end{bmatrix},$$

we get

$$M_1(t) = M_4(t) = \sum_{k=0}^{\infty}\frac{t^{2k}\tilde{v}^k}{(2k)!},\ M_2(t) = -\frac{1}{2}\sum_{k=0}^{\infty}\frac{t^{2k+1}\tilde{v}^{k+1}}{(2k+1)!},\ M_3(t) = -2\sum_{k=0}^{\infty}\frac{t^{2k+1}\tilde{v}^k}{(2k+1)!}.$$

If $\tilde{v} = v + b^2 \in \mathcal{S}_d^{+,*}$, $\sqrt{\tilde{v}}$ is well defined and we have $M_1(t) = M_4(t) = \cosh(t\sqrt{\tilde{v}})$, $M_2(t) = -\frac{1}{2}\sqrt{\tilde{v}}\sinh(t\sqrt{\tilde{v}})$ and $M_3(t) = -2(\sqrt{\tilde{v}})^{-1}\sinh(t\sqrt{\tilde{v}})$. Now, we define

$$V(t) = M_4(t) - \frac{1}{2}M_3(t)\tilde{w} = \left(\sum_{k=0}^{\infty} t^{2k+1}\frac{\tilde{v}^k}{(2k+1)!}\right)\tilde{w} + \sum_{k=0}^{\infty} t^{2k}\frac{\tilde{v}^k}{(2k)!}.$$

Since $V'(t) = -2M_2(t) + M_1(t)\tilde{w}$, we obtain that

$$\tilde{\gamma}(t) = -\frac{1}{2}V'(t)V(t)^{-1} \text{ and thus } \gamma(t) = -\frac{1}{2}\left(V'(t)V(t)^{-1} + b\right).$$

Last, we have $\beta'(t) = -\frac{1}{2}\alpha\,\mathrm{Tr}[V'(t)V(t)^{-1}] - \frac{1}{2}\alpha\,\mathrm{Tr}[b]$ and we obtain that

$$\beta(t) = -\frac{1}{2}\alpha\log(\det[V(t)]) - \frac{1}{2}\alpha\,\mathrm{Tr}[b]t,$$

since $\frac{d\det[V(t)]}{dt} = \det[V(t)]\,\mathrm{Tr}[V'(t)V(t)^{-1}]$.

It remains to show that we indeed have (54) for v and w satisfying (53). We define $\mathcal{E}_t = \frac{\exp\left(\beta(T-t)+\mathrm{Tr}[\gamma(T-t)X_t]+\mathrm{Tr}[-\frac{v}{2}R_t]\right)}{\exp(\beta(T)+\mathrm{Tr}[\gamma(T)x])}$. By Itô's formula, we have

$$\frac{d\mathcal{E}_t}{\mathcal{E}_t} = \mathrm{Tr}[\gamma(T-t)(\sqrt{X_t}dW_t + dW_t^\top\sqrt{X_t})].$$

This is a positive local martingale and thus a supermartingale which gives $\mathbb{E}[\mathcal{E}_T] \leqslant 1$, and we want to prove that this is a martingale. To do so, we use the argument presented by Rydberg in [34]. For $L > 0$, we define

$$\tau_L = \inf\{t \geqslant 0, \mathrm{Tr}[X_t] \geqslant L\},$$

and $\pi_L(x) = x\mathbf{1}_{\mathrm{Tr}[x]\leqslant L} + \frac{L}{\mathrm{Tr}[x]}x\mathbf{1}_{\mathrm{Tr}[x]>L}$ for $x \in \mathcal{S}_d^+$. We consider $(\mathcal{E}_t^L, t \in [0,T])$ the solution of

$$d\mathcal{E}_t^L = \mathcal{E}_t^L\,\mathrm{Tr}[\gamma(T-t)(\sqrt{\pi_L(X_t)}dW_t + dW_t^\top\sqrt{\pi_L(X_t)})],\ \mathcal{E}_0^L = 1.$$

We clearly have $\mathbb{E}[\mathcal{E}_T^L]=1$. Besides, under $\mathbb{P}^L$ given by $\frac{d\mathbb{P}^L}{d\mathbb{P}}\big|_{\mathcal{F}_T}=\mathcal{E}_T^L$, the process

$$dW_t^L=dW_t-2\sqrt{\pi_L(X_t)}\gamma(T-t)dt,\ t\in[0,T],$$

is a matrix Brownian motion. Since $\mathcal{E}_t=\mathcal{E}_t^L$ for $t\leqslant\tau_L$, we have $\mathbb{E}[\mathcal{E}_T]=\mathbb{E}[\mathcal{E}_T^L\mathbf{1}_{\tau_L>T}]+\mathbb{E}[\mathcal{E}_T\mathbf{1}_{\tau_L\leqslant T}]$. By Lebesgue's theorem, we get $\mathbb{E}[\mathcal{E}_T\mathbf{1}_{\tau_L\leqslant T}]\underset{L\to+\infty}{\to}0$. On the other hand, $\mathbb{E}[\mathcal{E}_T^L\mathbf{1}_{\tau_L>T}]=\mathbb{P}^L(\tau_L>T)$. Let us consider the Wishart process $\tilde{X}$ starting from x such that

$$d\tilde{X}_t=\Big[\alpha I_d+(b+2\gamma(T-t))\tilde{X}_t+\tilde{X}_t(b+2\gamma(T-t))\Big]dt+\sqrt{\tilde{X}_t}dW_t+dW_t^\top\sqrt{\tilde{X}_t}.$$

We also define $\tilde{\tau}_L=\inf\{t\in[0,T],\mathrm{Tr}[\tilde{X}_t]\geqslant L\}$ with convention $\inf\emptyset=+\infty$. The process $\tilde{X}$ solves the same SDE on $[0,\tilde{\tau}_L\wedge T]$ under $\mathbb{P}$ as X on $[0,\tau_L\wedge T]$ under $\mathbb{P}^L$. We therefore have

$$\mathbb{P}^L(\tau_L>T)=\mathbb{P}(\tilde{\tau}_L>T)\underset{L\to+\infty}{\to}1,$$

which finally gives $\mathbb{E}[\mathcal{E}_T]=1$. □

Corollary 5.1. *Let $Y\sim WIS_d(y,\alpha,b,a)$ be a Wishart process with parameters $\alpha\geqslant d-1$, $y\in\mathcal{S}_d^+$, $a,b\in\mathcal{M}_d$ satisfying*

$$ba^\top a=a^\top ab^\top \text{ and } a \text{ invertible.} \tag{60}$$

Let $v,w\in\mathcal{S}_d$ be such that

$$\exists m\in\mathcal{S}_d,\ \frac{1}{2}awa^\top+m\in\mathcal{S}_d^+ \text{ and } \frac{ava^\top}{2}-ab^\top a^{-1}m-m(a^\top)^{-1}ba^\top-2m^2\in\mathcal{S}_d^+. \tag{61}$$

Then, we have

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\frac{1}{2}\mathrm{Tr}\left[wY_T+v\int_0^T Y_s ds\right]\right)\right]\\&=\frac{\exp\big(-\frac{\alpha}{2}\mathrm{Tr}[b]t\big)}{\det[V_{v,w}(t)]^{\frac{\alpha}{2}}}\exp\Big(-\frac{1}{2}\mathrm{Tr}\left[(V'_{v,w}(t)V_{v,w}(t)^{-1}+(a^\top)^{-1}ba^\top)(a^\top)^{-1}ya^{-1}\right]\Big),\end{aligned}$$

with $V_{v,w}(t)=\left(\sum_{k=0}^\infty t^{2k+1}\frac{\tilde{v}^k}{(2k)!}\right)\tilde{w}+\sum_{k=0}^\infty t^{2k}\frac{\tilde{v}^k}{(2k)!}$ and

$$\tilde{v}=ava^\top+(a^\top)^{-1}b^2a^\top,\quad \text{and}\quad \tilde{w}=awa^\top-(a^\top)^{-1}ba^\top.$$

Proof. We know that $Y\underset{law}{=}a^\top Xa$ with $x=(a^\top)^{-1}ya^{-1}$ and $X\sim WIS_d(x,\alpha,(a^\top)^{-1}ba^\top,I_d)$, see e.g. equation (13) in [1]. We notice that $(a^\top)^{-1}ba^\top=ab^\top a^{-1}\iff ba^\top a=a^\top ab^\top$ and thus $(a^\top)^{-1}ba^\top\in\mathcal{S}_d$. We have

$$\mathbb{E}\left[\exp\left(-\frac{1}{2}\mathrm{Tr}\left[wY_T+v\int_0^T Y_s ds\right]\right)\right]=\mathbb{E}\left[\exp\left(-\frac{1}{2}\mathrm{Tr}\left[awa^\top X_T+ava^\top\int_0^T X_s ds\right]\right)\right],$$

which gives the result by applying Proposition 5.1. □

By setting $\tilde{m} = a^{-1}m(a^\top)^{-1}$, the condition (61) is equivalent to the existence of $\tilde{m} \in \mathcal{S}_d$, such that

$$\frac{1}{2}w + \tilde{m} \in \mathcal{S}_d^+ \text{ and } \frac{v}{2} - b^\top \tilde{m} - \tilde{m}b - 2\tilde{m}a^\top a\tilde{m} \in \mathcal{S}_d^+. \tag{62}$$

The case $m = 0$ gives back the finiteness of the Laplace transform when $v, w \in \mathcal{S}_d^+$. If we take $\tilde{m} = -w/2$, we get also the finiteness when

$$v + b^\top w + wb - wa^\top aw \in \mathcal{S}_d^+. \tag{63}$$

Another interesting choice is $m = -\frac{1}{2}(a^\top)^{-1}ba^\top$. We have $m \in \mathcal{S}_d$ from (60). This choice gives the finiteness of the Laplace transform when $v+b^\top(a^\top a)^{-1}b \in \mathcal{S}_d^+$ and $w-(a^\top a)^{-1}b \in \mathcal{S}_d^+$. Let us note that $\tilde{v} = a(v + b^\top(a^\top a)^{-1}b)a^\top$ so that the first condition is the same as $\tilde{v} \in \mathcal{S}_d^+$. Another interesting choice of m is given by the next remark.

Remark 5.1. *Proposition 5.1 extends the result of Gnoatto and Grasselli [17] to $\alpha \geqslant d-1$, and the sufficient condition* (61) *that ensures the finiteness of the Laplace transform is also less restrictive, which is crucial in our study especially in the nonergodic case. In particular, it does not assume a priori that $v+b^\top(a^\top a)^{-1}b \in \mathcal{S}_d^+$. We can recover the result of [17] as follows. Let us assume $v+b^\top(a^\top a)^{-1}b \in \mathcal{S}_d^+$ and take $m = -\frac{(a^\top)^{-1}ba^\top}{2} + \frac{1}{2}\sqrt{a(v + b^\top(a^\top a)^{-1}b)a^\top}$. We have $m \in \mathcal{S}_d$ from* (60) *and it satisfies $\frac{ava^\top}{2} - ab^\top a^{-1}m - m(a^\top)^{-1}ba^\top - 2m^2 = 0 \in \mathcal{S}_d^+$. Therefore,* (61) *holds if*

$$w - (a^\top a)^{-1}b + a^{-1}\sqrt{a(v + b^\top(a^\top a)^{-1}b)a^\top}(a^\top)^{-1} \in \mathcal{S}_d^+.$$

This is precisely the condition stated in [17].

Remark 5.2. *It is possible to get similarly the Laplace transform of $\left(Y_T, \int_0^T Y_s ds\right)$ when Y solves*

$$dY_t = \left[\overline{\alpha} + bY_t + Y_t b^\top\right] dt + \sqrt{Y}_t dW_t a + a^\top dW_t^\top \sqrt{Y}_t,\ Y_0 = y \in \mathcal{S}_d^+,$$

with a, b satisfying (60) *and $\overline{\alpha} - (d-1)a^\top a \in \mathcal{S}_d^+$. Again, equation (13) in [1] gives $Y \underset{law}{=} a^\top X a$, where*

$$dX_t = \left[\hat{\alpha} + \hat{b}X_t + X_t\hat{b}^\top\right] dt + \sqrt{X}_t dW_t + dW_t^\top \sqrt{X}_t,\ X_0 = x,$$

with $x = (a^\top)^{-1}ya^{-1} \in \mathcal{S}_d$, $\hat{b} = (a^\top)^{-1}ba^\top \in \mathcal{S}_d$ and $\hat{\alpha} = (a^\top)^{-1}\overline{\alpha}a^{-1} \in \mathcal{S}_d$. Repeating the proof of Proposition 5.1, we observe that the Riccati equation (58) *and equation* (57) *remain unchanged while* (59) *is replaced by*

$$\beta' = \mathrm{Tr}[\hat{\alpha}\gamma] = -\frac{1}{2}\mathrm{Tr}[\hat{\alpha}V'(t)V(t)^{-1}] - \frac{1}{2}\mathrm{Tr}[\hat{\alpha}\hat{b}].$$

Therefore, we deduce that under the same condition (61), *we have*

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\frac{1}{2}\mathrm{Tr}\left[wY_T + v\int_0^T Y_s ds\right]\right)\right]\\ &= \exp(\beta(T))\exp\Big(-\frac{1}{2}\mathrm{Tr}\left[(V'_{v,w}(t)V_{v,w}(t)^{-1} + (a^\top)^{-1}ba^\top)(a^\top)^{-1}ya^{-1}\right]\Big),\end{aligned}$$

with $\beta(t) = -\frac{1}{2}\int_0^t \mathrm{Tr}[(a^\top)^{-1}\overline{\alpha}a^{-1}V'_{v,w}(s)V_{v,w}(s)^{-1}]ds - \frac{t}{2}\mathrm{Tr}[\overline{\alpha}(a^\top a)^{-1}b]$ *and* $V_{v,w}(t)$ *defined as in Corollary 5.1. Thus, the formula is no longer totally explicit. In Gnoatto and Grasselli [17], the result is stated with* $\mathrm{Tr}[(a^\top)^{-1}\overline{\alpha}a^{-1}\log(V_{v,w}(t))]$ *instead of the first integral. However, this replacement does not seem clear to us unless* $V'_{v,w}(s)$ *and* $V_{v,w}(s)$ *commute for all* $s \geqslant 0$ *(this happens when the matrices* $\tilde{v}$ *and* $\tilde{w}$ *in* $V_{v,w}$ *commute) or* $\overline{\alpha} = \alpha a^\top a$ *by using the trace cyclic theorem.*

Corollary 5.2. *Let* $Y \sim WIS_d(y, \alpha, b, a)$ *be a Wishart process with parameters such that* $ba^\top a = a^\top ab^\top$ *and* a *invertible. Then,*

$$\forall u \in \mathcal{S}_d,\ \mathbb{E}\left[\exp\left(\int_0^T \mathrm{Tr}[u\sqrt{Y_s}dW_s a]ds - \frac{1}{2}\int_0^T \mathrm{Tr}[auY_s ua^\top]ds\right)\right] = 1.$$

Proof. We have $2\int_0^T \mathrm{Tr}[u\sqrt{Y_s}dW_s a]ds = \mathrm{Tr}[u(Y_T - y)] - \alpha T\,\mathrm{Tr}[ua^\top a] - \mathrm{Tr}\left[(ub + b^\top u)\int_0^T Y_s ds\right]$. We apply Corollary 5.1 with $w = -u$ and $v = ub + b^\top u + ua^\top au$. Therefore, (63) holds. We then have $\tilde{w} = -(aua^\top + (a^\top)^{-1}ba^\top)$ and $\tilde{v} = \tilde{w}^2$ and the result follows by simple calculations. □

5.2 Study of the MLE of b with the Laplace transform

We consider $\epsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$ a (deterministic) decreasing function such that $\lim_{t\to+\infty} \epsilon_t = 0$. From the definition of the MLE of b (18), we get that

$$\frac{1}{\epsilon_T}(\hat{b}_T - b) = \mathcal{L}^{-1}_{\epsilon_T^2 R_T}(\epsilon_T[X_T - x - \alpha T I_d - bR_T - R_T b]).$$

Thus, we want to calculate the Laplace transform of $(\epsilon_T[X_T - x - \alpha T I_d - bR_T - R_T b], \epsilon_T^2 R_T)$ in order to study the convergence of $\frac{1}{\epsilon_T}(\hat{b}_T - b)$. For $\lambda_1, \lambda_2 \in \mathcal{S}_d$, we define

$$\mathcal{E}(T, \lambda_1, \lambda_2) := \mathbb{E}_\theta\Big[\exp\Big(-\epsilon_T\,\mathrm{Tr}[\lambda_2(X_T - x - \alpha T I_d - bR_T - R_T b)] - \epsilon_T^2\,\mathrm{Tr}[\lambda_1 R_T]\Big)\Big] \quad (64)$$

$$= \exp\left(\epsilon_T\,\mathrm{Tr}[\lambda_2(x + \alpha T I_d)]\right)\mathbb{E}_\theta\Big[\exp\Big(-\mathrm{Tr}[\epsilon_T\lambda_2 X_T] - \mathrm{Tr}[(\epsilon_T^2\lambda_1 - \epsilon_T(\lambda_2 b + b\lambda_2))R_T]\Big)\Big]. \quad (65)$$

We now consider $\lambda_1, \lambda_2 \in \mathcal{S}_d$ such that

$$\lambda_1 - 2\lambda_2^2 \in \mathcal{S}_d^{+,*}. \quad (66)$$

We define

$$v_T = 2\lambda_1\epsilon_T^2 - 2(b\lambda_2 + \lambda_2 b)\epsilon_T, \quad \tilde{v}_T = v_T + b^2, \quad w_T = 2\lambda_2\epsilon_T, \quad \tilde{w}_T = w_T - b, \quad (67)$$

and have $v_T + bw_T + w_T b - w_T^2 = \epsilon_T^2(2\lambda_1 - 4\lambda_2^2) \in \mathcal{S}_d^{+,*}$. Thus, by applying Proposition 5.1 with $m = -\epsilon_T\lambda_2$, we get that $\mathcal{E}(T, \lambda_1, \lambda_2)$ is finite and given by

$$\begin{aligned}\mathcal{E}(T, \lambda_1, \lambda_2) =& \frac{\exp\left(-\frac{\alpha}{2}\mathrm{Tr}[b]T\right)}{\det[V_{v_T,w_T}(T)]^{\frac{\alpha}{2}}}\exp\Big(-\frac{1}{2}\mathrm{Tr}\left[(V'_{v_T,w_T}(T)V_{v_T,w_T}(T)^{-1} + b)x\right]\Big)\\ &\times\exp\Big(\epsilon_T\,\mathrm{Tr}\left[\lambda_2(x + \alpha T I_d)\right]\Big)\end{aligned} \quad (68)$$

with

$$\begin{aligned} V_{v_T,w_T}(T) =& (\sqrt{\tilde{v}_T})^{-1}\sinh(\sqrt{\tilde{v}_T}T)\tilde{w}_T + \cosh(\sqrt{\tilde{v}_T}T) \\ V'_{v_T,w_T}(T) =& \cosh(\sqrt{\tilde{v}_T}T)\tilde{w}_T + \sinh(\sqrt{\tilde{v}_T}T)\sqrt{\tilde{v}_T}. \end{aligned}$$

Besides, we have $\tilde{v}_T = (b - 2\epsilon_T\lambda_2)^2 + \epsilon_T^2(2\lambda_1 - 4\lambda_2^2) \in \mathcal{S}_d^{+,*}$.

When $-b \in \mathcal{S}_d^{+,*}$ and $\epsilon_T = 1/\sqrt{T}$, we can make explicit calculations and get

$$\lim_{T\to+\infty} \mathcal{E}(T,\lambda_1,\lambda_2) = \exp(-\operatorname{Tr}[\lambda_1\overline{R}_\infty] - \operatorname{Tr}[2\lambda_2^2\overline{R}_\infty]),$$

which gives another mean to prove Theorem 2.4. Here, we prove Theorem 3.4.

Proof of Theorem 3.4. Here, we focus on the case $b = b_0 I_d$ with $b_0 > 0$ and set $\epsilon_T = e^{-b_0T}$. Since the square root function is analytic on the set of positive definite matrices (see e.g. [33], p. 134) we get that

$$\sqrt{\tilde{v}_T} = b_0I_d - 2\epsilon_T\lambda_2 + \frac{\epsilon_T^2}{b_0}(\lambda_1 - 2\lambda_2^2) + O(\epsilon_T^3),$$

since the squares of each sides coincides up to a $O(\epsilon_T^3)$ term. We observe that $\tilde{w}_T = 2\epsilon_T\lambda_2 - b_0I_d$, and thus $\sqrt{\tilde{v}_T} + \tilde{w}_T = \frac{\epsilon_T^2}{b_0}(\lambda_1 - 2\lambda_2^2) + O(\epsilon_T^3)$.

We now write

$$\begin{aligned} V_{v_T,w_T}(T) =& (\sqrt{\tilde{v}_T})^{-1}\left[\frac{1}{2}\exp(\sqrt{\tilde{v}_T}T)(\sqrt{\tilde{v}_T} + \tilde{w}_T) + \frac{1}{2}\exp(-\sqrt{\tilde{v}_T}T)(\sqrt{\tilde{v}_T} - \tilde{w}_T)\right] \\ V'_{v_T,w_T}(T) =& \frac{1}{2}\exp(\sqrt{\tilde{v}_T}T)(\sqrt{\tilde{v}_T} + \tilde{w}_T) + \frac{1}{2}\exp(-\sqrt{\tilde{v}_T}T)(\tilde{w}_T - \sqrt{\tilde{v}_T}). \end{aligned}$$

Since $\epsilon_T\exp(\sqrt{\tilde{v}_T}T) \underset{T\to+\infty}{\to} I_d$, we get $\frac{1}{\epsilon_T}V_{v_T,w_T}(T) \underset{T\to+\infty}{\to} \frac{1}{b_0}\left[\frac{1}{2b_0}(\lambda_1 - 2\lambda_2^2) + b_0I_d\right]$ and $\frac{1}{\epsilon_T}V'_{v_T,w_T}(T) \underset{T\to+\infty}{\to} \frac{1}{2b_0}(\lambda_1 - 2\lambda_2^2) - b_0I_d$. This yields to

$$V'_{v_T,w_T}(T)V_{v_T,w_T}(T)^{-1} + b_0I_d \underset{T\to+\infty}{\longrightarrow} (\lambda_1 - 2\lambda_2^2)\left(\frac{1}{2b_0}(\lambda_1 - 2\lambda_2^2) + b_0I_d\right)^{-1}.$$

We also have $\frac{\exp\left(-\frac{\alpha}{2}\operatorname{Tr}[b_0I_d]T\right)}{\det[V_{v_T,w_T}(T)]^{\frac{\alpha}{2}}} = \frac{1}{\det[\epsilon_T^{-1}V_{v_T,w_T}(T)]^{\frac{\alpha}{2}}} \underset{T\to+\infty}{\to} \frac{1}{\det\left[\frac{1}{b_0}\left[\frac{1}{2b_0}(\lambda_1-2\lambda_2^2)+b_0I_d\right]\right]}$, and therefore

$$\lim_{T\to+\infty} \mathcal{E}(T,\lambda_1,\lambda_2) = \frac{\exp\left(-\frac{1}{2b_0}\operatorname{Tr}\left[(\lambda_1 - 2\lambda_2^2)\left(\frac{1}{2b_0^2}(\lambda_1 - 2\lambda_2^2) + I_d\right)^{-1}x\right]\right)}{\det\left[\frac{1}{2b_0^2}(\lambda_1 - 2\lambda_2^2) + I_d\right]}. \tag{69}$$

We now want to identify the limit. We know that $X \sim WIS_d\left(\frac{x}{2b_0}, \alpha, 0, I_d; \frac{1}{4b_0^2}\right)$ has the following Laplace transform

$$u \in \mathcal{S}_d^+, \mathbb{E}[\exp(-\operatorname{Tr}[uX])] = \frac{\exp\left(-\operatorname{Tr}\left[u(I_d + \frac{1}{2b_0^2}u)^{-1}\frac{x}{2b_0}\right]\right)}{\det[I_d + \frac{1}{2b_0^2}u]}.$$

Let $\tilde{\mathbf{G}}$ denote a d-square matrix independent from X, whose entries are independent and follow a standard Normal distribution. By Lemma B.4, we have

$$\mathbb{E}[\exp(-\operatorname{Tr}[\lambda_1 X + \lambda_2(\sqrt{X}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}\sqrt{X})])] = \mathbb{E}[\exp(-\operatorname{Tr}[(\lambda_1 - 2\lambda_2^2)X])].$$

Thus, (69) shows the convergence in law of $\left(\epsilon_T(X_T - x - \alpha T I_d - bR_T - R_T b), \epsilon_T^2 R_T\right)$ to $\left(X, \sqrt{X}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}\sqrt{X}\right)$ under $\mathbb{P}_\theta$, which gives the claim of Theorem 3.4. $\square$

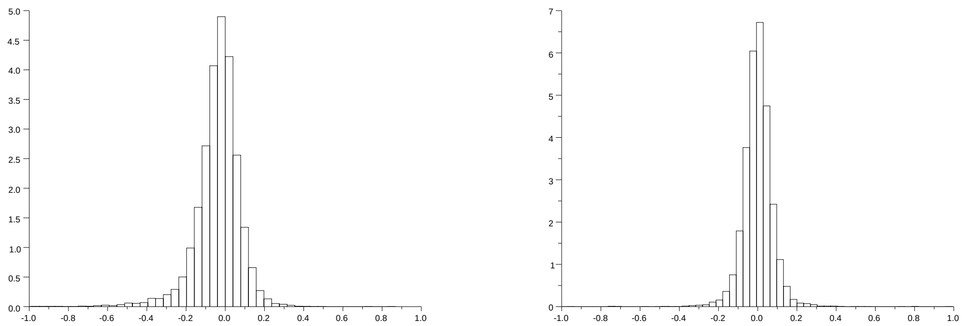
6 Numerical Study

In this section, we test the convergence of the MLE given by (15) and (18). To do so, we consider a given large value of T and simulate the Wishart process exactly on the regular time grid $t_i = \frac{iT}{N}$, $i = 0, \cdots, N$. This can be done by using the method presented in Ahdida and Alfonsi [1], see also Alfonsi [3]. We take N sufficiently large and approximate the integrals R_T and Q_T^{-1} applying the trapezoidal rule along this time grid. Thus, we will use the estimator with the exact value of X_T and these approximated values of R_T and Q_T^{-1}.

This section has three goals. First, we check numerically the convergence results that we have obtained. Second, we investigate numerically the convergence of the MLE in some nonergodic cases, where no theoretical result of convergence have been found. Last, we test the estimation of the parameters of a full Wishart process (1). To do so, we estimate first a with the quadratic variation and then the parameters α and b by using the MLE (15) on the process $(a^\top)^{-1} X a^{-1}$.

6.1 Numerical validation of the convergence results

Using the method mentioned above, we have checked the convergence results obtained in this paper. Namely, we sample $M = 10000$ independent paths of X in order to draw an histogram of the properly rescaled value of $\hat{b}_{i,j} - b_{i,j}$ or $\hat{\alpha} - \alpha$. We do not reproduce all these graphics here, and present for example in Figure 1 an illustration of the convergence given by Theorem 3.4.



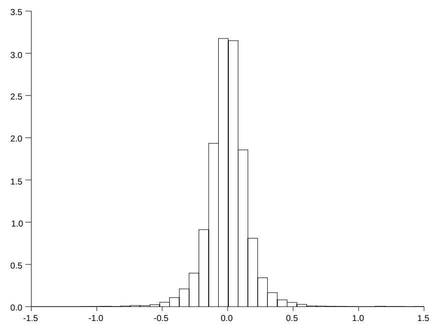
(a) Limit law of $\exp(0.05T)(b - \hat{b}_T)_{1,1}$.

(b) Limit law of $\exp(0.05T)(b - \hat{b}_T)_{1,2}$.

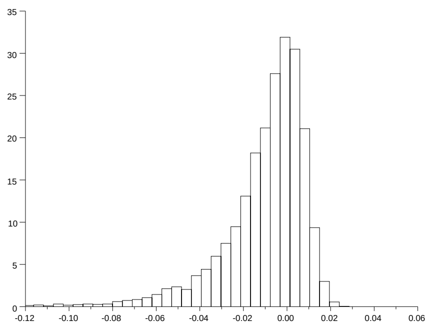
Figure 1: Asymptotic law of the error for the estimation of $\theta = b$ with for: $x = \left(\begin{smallmatrix} 0.5 & 0.1 \\ 0.1 & 0.3 \end{smallmatrix}\right)$, $T = 100$, $N = 10000$, $\alpha = 4.5$ and $b = 0.05 I_d$.

6.2 Experimental convergence in a nonergodic case

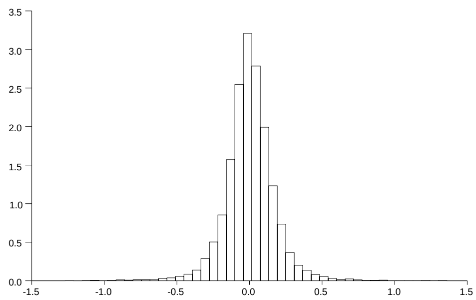
In this paragraph, we try to guess the asymptotic behavior of the MLE in a nonergodic case, where no theoretical convergence result is known. Namely, we observe in Figure 2 the asymptotic estimation error, when $b = \mathrm{diag}(0.1, 0.005)$ is diagonal with positive and distinct terms on its diagonal and when we use the estimator (18). As one might have guess, the convergence of the diagonal terms seems to be with an exponential rate, with the exponential speed corresponding to its value. Namely, $\hat{b}_{11}$ seems to converge to b_{11} with a speed of $\exp(0.1T)$ while $\hat{b}_{22}$ seems to converge to b_{22} with a speed of $\exp(0.005T)$. More interesting is the antidiagonal term. One could have imagine that the convergence rate is the slowest of these two rates. Instead, on our experiment, the convergence of $\hat{b}_{12}$ towards b_{12} seems to happen with the rate $\exp(0.1T)$. We have observed the same behaviour for other parameter values. Of course, it would be hasty to draw a global conclusion from few particular experiments. However, it is interesting to note that these numerical tests are a way to guess or check the convergence rate of the MLE.



(a) Limit law of $\exp(0.1T)(b - \hat{b}_T)_{1,1}$.



(b) Limit law of $\exp(0.005T)(b - \hat{b}_T)_{2,2}$.



(c) Limit law of $\exp(0.1T)(b - \hat{b}_T)_{1,2}$.

Figure 2: Asymptotic law of the error for the estimation of $\theta = b$ with $x = \left(\begin{smallmatrix} 0.3 & 0.1 \\ 0.1 & 0.2 \end{smallmatrix}\right)$, $T = 100$, $N = 10000$, $\alpha = 3.5$ and $b = \mathrm{diag}(0.1, 0.005)$.

6.3 Estimation of the whole Wishart process

In this last part of the numerical study, we perform the estimation of all the parameters of the Wishart process (1). We consider a case where a is upper triangular and $(a^\top)^{-1}ba^\top$ is symmetric. We proceed as follows. First, we sample exactly a discrete path $(X_{iT/N}, 0 \leqslant i \leqslant N)$.

Then, we estimate the matrix $a^\top a$ by using (3), where the quadratic variations are replaced by their classical approximations and the integrals are replaced by the trapezoidal rule. By a Cholesky decomposition we get then an estimator $\hat{a}$ of a. Then, we use the MLE (15) on the path $((\hat{a}^\top)^{-1}X_{iT/N}\hat{a}^\top, 0 \leqslant i \leqslant N)$. This gives an estimator of α and $(a^\top)^{-1}ba^\top$, and therefore an estimator of b. As a comparison, we also calculate similarly the estimator of α and b when a is known and has not to be estimated. To draw histograms or calculate empirical expectations, we run $M = 10000$ independent paths of X.

We consider a sufficiently large value of T and are interested in looking at the convergence with respect to N. First, we plot the the error on the estimator of a with respect to the number of time step in Log-Log scale. We observe that the convergence to zero takes place with experimental rate close to 1/2. This is in line with the general results on the estimation of the diffusion coefficient, see Dohnal [12] and Genon-Catalot and Jacod [15]. Then, we focus on the influence of the discretization and the unknown parameter a on the convergence of the MLE of b and α. In Table 1, we give in function of N the Mean Squared Error $\mathrm{MSE}(\widehat{\theta}^N|\theta) = \mathbb{E}[|\widehat{\theta}^N - \theta|^2]$ of the estimator $\widehat{\theta}^N$, with $\theta = (b, \alpha)$. It is estimated with the empirical expectation. First, we observe that the convergence of the estimator of α is roughly the same whether we know a or not. This is expected since the estimation of α does not depend on the estimation of a. Instead, the bias on b is much higher when a is estimated than when a is known. However, it decreases also faster at an experimental order of 0.7 while the bias when a is known decreases at an experimental order of 0.45. This latter rate is in line with the rate of 1/2 obtained in dimension 1 by Ben Alaya and Kebaier [5]. In our case, it seems that the influence of the estimation of a vanishes around $N = 5000$. Last, we have plotted in Figure 4 the limit law of the estimator $\sqrt{T}(\widehat{\theta}^N - \theta)$ with $N = 10000$.

This short numerical study shows that the estimator obtained by discretizing the continuous time estimator is efficient in practice. Of course, it would be nice to obtain general convergence results in function of T and N, but we leave this for further research.

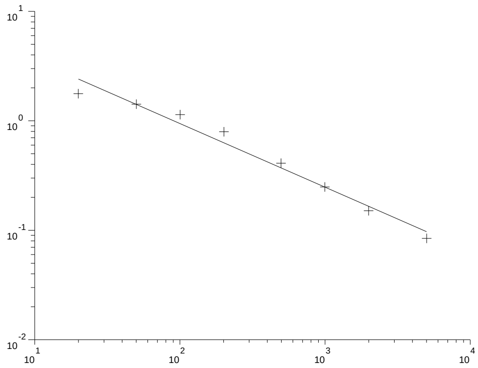
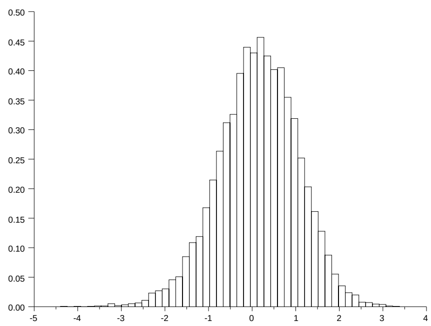


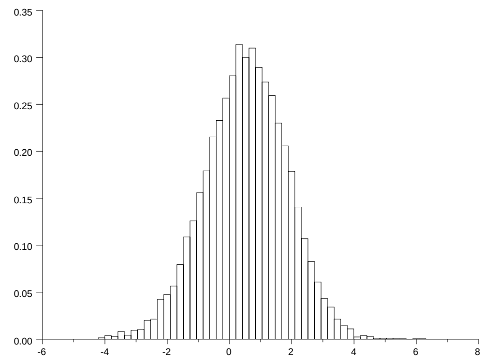
Figure 3: Log-Log representation of the empirical expectation of $\mathbb{E}[\mathrm{Tr}[(a - \widehat{a}^N)^2]]^{1/2}$ for $x = \begin{pmatrix} 0.8 & 0.5 \\ 0.5 & 1 \end{pmatrix}$, $T = 100$, $a = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, $\alpha = 4.5$, $b = \begin{pmatrix} -1 & 0.2 \\ 2 & -2 \end{pmatrix}$, where the line is the simple linear regression i.e. $\log(\mathbb{E}[\mathrm{Tr}[(a - \widehat{a}^N)^2]]^{1/2}) \approx 2.62 - 0.58 \log(N)$.

Number of time steps		20	50	100	200	500	1000	2000	5000
$\mathbb{E}[\mathrm{Tr}[(a-\widehat{a}^N)^2]]^{1/2}$		1.7671	1.4311	1.1487	0.7913	0.4107	0.2472	0.1514	0.0846
$\mathrm{MSE}(\widehat{b}^N_{1,1}\|b_{1,1})$	$\widehat{a}=a$	0.0745	0.0338	0.0181	0.0115	0.0082	0.0069	0.0061	0.0058
	$\widehat{a}=\widehat{a}^N$	0.7636	0.5266	0.3489	0.1891	0.0624	0.0273	0.0142	0.0085
$\mathrm{MSE}(\widehat{b}^N_{2,2}\|b_{2,2})$	$\widehat{a}=a$	0.2554	0.1310	0.0664	0.0372	0.0231	0.0176	0.0153	0.0139
	$\widehat{a}=\widehat{a}^N$	3.4085	2.8722	2.1159	1.1995	0.3600	0.1264	0.0480	0.0201
$\mathrm{MSE}(\widehat{b}^N_{1,2}\|b_{1,2})$	$\widehat{a}=a$	0.0075	0.0033	0.0017	0.0011	0.0008	0.0008	0.0007	0.0007
	$\widehat{a}=\widehat{a}^N$	0.0442	0.0568	0.0596	0.0352	0.0148	0.0075	0.0039	0.0019
$\mathrm{MSE}(\widehat{\alpha}^N\|\alpha)$	$\widehat{a}=a$	0.8448	0.3579	0.1993	0.1151	0.0614	0.0416	0.0308	0.0230
	$\widehat{a}=\widehat{a}^N$	0.8267	0.3496	0.1895	0.1095	0.0617	0.0410	0.0311	0.0234

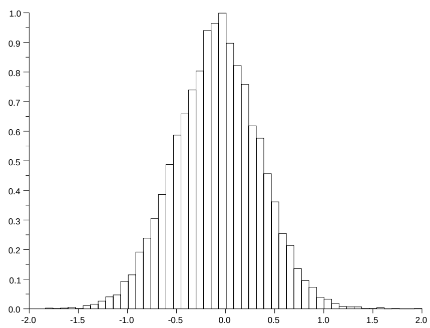
Table 1: Mean Squared Error for the estimation of $\theta = (b, \alpha)$ with respect to N. Same parameters as Figure 3.



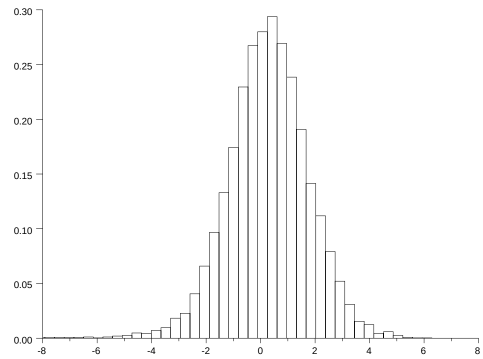
(a) Limit law of $\sqrt{T}(b - \hat{b}^N_T)_{1,1}$.



(b) Limit law of $\sqrt{T}(b - \hat{b}^N_T)_{2,2}$.



(c) Limit law of $\sqrt{T}(b - \hat{b}^N_T)_{1,2}$.



(d) Limit law of $\sqrt{T}(\alpha - \hat{\alpha}^N_T)$.

Figure 4: Asymptotic laws of the error for the estimation of $\theta = (b, \alpha)$ for $\widehat{a} = \widehat{a}^N$, $N = 10000$, same parameters as Figure 3.

A Proof of Proposition 1.1

We denote $b^s = (b + b^\top)/2$ (resp. $b^a = (b - b^\top)/2$) the symmetric (resp. antisymmetric) part of b. We have

$$\int_0^T \mathrm{Tr}[(\sqrt{X_s})^{-1} dW_s] = \frac{1}{2}\log\left(\frac{\det[X_T]}{\det[x]}\right) - \mathrm{Tr}[b]T - \frac{1}{2}\int_0^T (\alpha - 1 - d)\,\mathrm{Tr}[X_s^{-1}]ds,$$

$$\begin{aligned}\int_0^T \mathrm{Tr}[b^s\sqrt{X_s}dW_s] &= \frac{1}{2}\int_0^T \mathrm{Tr}[b^s(\sqrt{X_s}dW_s + dW_s^\top\sqrt{X_s})] \\ &= \frac{\mathrm{Tr}[b^s X_T] - \mathrm{Tr}[b^s x]}{2} - \frac{\alpha T}{2}\mathrm{Tr}[b^s] - \frac{1}{2}\int_0^T \mathrm{Tr}[b^s(bX_t + X_t b^\top)]dt.\end{aligned}$$

Thus, the only part to calculate is $\mathbb{E}\left[\exp(\int_0^T \mathrm{Tr}[-b^a\sqrt{X_s}dW_s])\middle|\mathcal{F}_T^X\right]$, and we set $M_t^a = \int_0^t \mathrm{Tr}[b^a\sqrt{X_s}dW_s]$. We now observe that $\langle \mathrm{Tr}[A_s dW_s], \mathrm{Tr}[B_s dW_s]\rangle = \mathrm{Tr}[A_s B_s^\top]ds$ and are looking for the process Γ that takes values in $\mathcal{S}_d$ and minimizes

$$\langle dM_t^a - \mathrm{Tr}[\Gamma_t(dX_t - (\alpha I_d + bX_t + X_t b^\top)dt)]\rangle = \left\{-\mathrm{Tr}[b^a X_t b^a] + 2\,\mathrm{Tr}[\Gamma_t(X_t b^a - b^a X_t)] + 4\,\mathrm{Tr}[\Gamma_t^2 X_t]\right\}dt.$$

We obtain that $2(X_t b^a - b^a X_t) + 4(X_t\Gamma_t + \Gamma_t X_t) = 0$ and thus

$$\Gamma_t = \mathcal{L}_{X_t}^{-1}\left(\frac{1}{2}(b^a X_t - X_t b^a)\right).$$

It satisfies $\mathrm{Tr}[\Gamma_t(X_t b^a - b^a X_t)] = -2\,\mathrm{Tr}[\Gamma_t(\Gamma_t X_t + X_t\Gamma_t)] = -4\,\mathrm{Tr}[\Gamma_t^2 X_t]$. By construction, we have $\langle dM_t^a - \mathrm{Tr}[\Gamma_t(dX_t - (\alpha I_d + bX_t + X_t b^\top)dt)], \mathrm{Tr}[\tilde{\Gamma}(dX_t - (\alpha I_d + bX_t + X_t b^\top)dt)]\rangle = 0$ for any $\tilde{\Gamma} \in \mathcal{S}_d$. Thus, there exists a Brownian motion β independent of X such that $dM_t^a - \mathrm{Tr}[\Gamma_t(dX_t - (\alpha I_d + bX_t + X_t b^\top)dt)] = \sqrt{-\mathrm{Tr}[b^a X_t b^a] - \mathrm{Tr}[\Gamma_t(b^a X_t - X_t b^a)]}d\beta_t$. In fact, both processes $(X_t, \int_0^t \sqrt{-\mathrm{Tr}[b^a X_s b^a] - \mathrm{Tr}[\Gamma_t(b^a X_s - X_s b^a)]}d\beta_s)$ and

$$(X_t, M_t^a - \int_0^t \mathrm{Tr}[\Gamma_s(\sqrt{X_s}dW_s + dW_s^\top\sqrt{X_s})])$$

solve the same martingale problem for which uniqueness holds. Therefore, we have

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\int_0^T \mathrm{Tr}[b^a\sqrt{X_t}dW_t]\right)\middle|\mathcal{F}_T^X\right] \\ &= \exp\left(\int_0^T \mathrm{Tr}\left[\Gamma_t((\alpha I_d + bX_t + X_t b^\top)dt - dX_t)\right] - \frac{1}{2}\int_0^T \mathrm{Tr}[b^a X_t b^a] + \mathrm{Tr}[\Gamma_t(b^a X_t - X_t b^a)]dt\right) \\ &= \exp\left(-\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(\frac{1}{2}(b^a X_t - X_t b^a)\right)dX_t\right] - \int_0^T \frac{1}{2}\mathrm{Tr}[b^a X_t b^a]dt\right. \\ &\quad \left.+\frac{1}{2}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(\frac{1}{2}(b^a X_t - X_t b^a)\right)(b^a X_t - X_t b^a)\right]dt + \int_0^T \mathrm{Tr}[b^a X_t b^s]dt\right),\end{aligned}$$

since $\mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(\frac{1}{2}(b^a X_t - X_t b^a)\right)\right] = \frac{1}{2}\mathrm{Tr}[X_t^{-1}(b^a X_t - X_t b^a)] = 0$ by Lemma B.1 and

$$\mathrm{Tr}[\Gamma_t(b^s X_t + X_t b^s)] = \mathrm{Tr}[b^s(\Gamma_t X_t + X_t\Gamma_t)] = \frac{1}{2}\mathrm{Tr}\left[b^s(b^a X_t - X_t b^a)\right] = \mathrm{Tr}[b^a X_t b^s].$$

Using (6) and the previous calculations, we obtain

$$L_T^{\theta,\theta_0} = \exp\Big(\frac{\alpha-\alpha_0}{4}\log\left(\frac{\det[X_T]}{\det[x]}\right) + \frac{\mathrm{Tr}[b^s X_T]-\mathrm{Tr}[b^s x]}{2} - \frac{1}{2}\int_0^T \mathrm{Tr}[(b^s)^2 X_s]ds$$
$$- \int_0^T \mathrm{Tr}[b^a X_s b^s]ds - \frac{\alpha-\alpha_0}{4}\big(\frac{\alpha+\alpha_0}{2}-1-d\big)\int_0^T \mathrm{Tr}[X_s^{-1}]ds - \frac{\alpha T}{2}\mathrm{Tr}[b]$$
$$+\frac{1}{2}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(b^a X_t - X_t b^a\right)dX_t\right] - \frac{1}{4}\int_0^T \mathrm{Tr}\left[\mathcal{L}_{X_t}^{-1}\left(b^a X_t - X_t b^a\right)\left(b^a X_t - X_t b^a\right)\right]dt\Big).$$

Last, we use $\mathcal{L}_{X_t}^{-1}\left(b^s X_t + X_t b^s\right) = b^s$ and $\mathrm{Tr}[\mathcal{L}_{X_t}^{-1}\left(b^a X_t - X_t b^a\right)\left(b^s X_t + X_t b^s\right)] = 2\,\mathrm{Tr}[b^a X_t b^s]$ to obtain (7).

B Technical lemmas

Lemma B.1. *For $X \in \mathcal{S}_d^{+,*}$ and $a \geqslant 0$, let $\mathcal{L}_{X,a}$ and $\mathcal{L}_X = \mathcal{L}_{X,0}$ be the linear applications defined by (12) on $\mathcal{S}_d$. If $a\,\mathrm{Tr}[X^{-1}] \neq 1$, then $\mathcal{L}_{X,a}$ is invertible and we have $\mathrm{Tr}[\mathcal{L}_{X,a}^{-1}(Y)] = \frac{\mathrm{Tr}[X^{-1}Y]}{2(1-a\,\mathrm{Tr}[X^{-1}])}$. Besides, the map $(X,Y,a) \mapsto \mathcal{L}_{X,a}^{-1}(Y)$ is continuous on $\{(X,Y,a) \in \mathcal{S}_d^{+,*} \times \mathcal{S}_d \times \mathbb{R}_+,\ a\,\mathrm{Tr}[X^{-1}] \neq 1\}$.*

Proof. The invertibility of $\mathcal{L}_{X,a}$ is equivalent to its one-to-one property. Since $X \in \mathcal{S}_d^{+,*}$, there exists an orthogonal matrix O_X and a diagonal matrix D_X with positive elements such that $X = O_X D_X O_X^\top$. We get

$$\begin{aligned} Y \in \ker(\mathcal{L}_{X,a}) &\iff O_X D_X O_X^\top Y + Y O_X D_X O_X^\top = 2a\,\mathrm{Tr}[Y] I_d \\ &\iff D_X(O_X^\top Y O_X) = 2a\,\mathrm{Tr}[Y] I_d - (O_X^\top Y O_X) D_X. \end{aligned} \tag{70}$$

Since D_X is diagonal, we obtain for $1 \leqslant i,k \leqslant d$, $\left((O_X^\top Y O_X)D_X\right)_{i,k} = (O_X^\top Y O_X)_{i,k}(D_X)_{k,k}$ and $\left(D_X(O_X^\top Y O_X)\right)_{i,k} = (D_X)_{i,i}(O_X^\top Y O_X)_{i,k}$. For $k \neq i$, (70) gives $(O_X^\top Y O_X)_{i,k} = 0$. For $k = i$, we get $(O_X^\top Y O_X)_{i,i}(D_X)_{i,i} = a\,\mathrm{Tr}[Y]$ and therefore

$$\mathrm{Tr}[Y] = \mathrm{Tr}[O_X^\top Y O_X] = \mathrm{Tr}[Y] a \sum_{i=1}^d \frac{1}{(D_X)_{i,i}} = \mathrm{Tr}[Y] a\,\mathrm{Tr}[X^{-1}].$$

Since $a\,\mathrm{Tr}[X^{-1}] \neq 1$, we obtain $\mathrm{Tr}[Y] = 0$ and then $(O_X^\top Y O_X)_{i,i} = 0$, which gives $Y = 0$ and the invertibility of $\mathcal{L}_{X,a}$. Let $c = \mathcal{L}_{X,a}^{-1}(Y)$. We have $c + X^{-1}cX - 2a\,\mathrm{Tr}[c]X^{-1} = X^{-1}Y$, which gives $2(1-a\,\mathrm{Tr}[X^{-1}])\,\mathrm{Tr}[c] = \mathrm{Tr}[X^{-1}Y]$. Last, the continuity property is obvious since $(X,a) \mapsto \mathcal{L}_{X,a}$ is continuous and $\mathcal{L} \mapsto \mathcal{L}^{-1}$ is continuous on $\{\mathcal{L} : \mathcal{S}_d \to \mathcal{S}_d$ linear and invertible$\}$. □

Lemma B.2. *For $X \in \mathcal{S}_d^{+,*}$, $\mathcal{L}_X$ is self-adjoint and positive definite:*

$$\mathrm{Tr}[\mathcal{L}_X(Y)Y] \geq 2\underline{\lambda}(X)\,\mathrm{Tr}[Y^2],$$

where $\underline{\lambda}(X) > 0$ is the lowest eigenvalue of X. Besides, for $a < 1/\mathrm{Tr}[X^{-1}]$, $\mathcal{L}_{X,a}$ is self-adjoint and positive definite.

Proof. For $Y, Z \in \mathcal{S}_d$, we have $\mathrm{Tr}[\mathcal{L}_X(Y)Z] = \mathrm{Tr}[(XY+YX)Z] = \mathrm{Tr}[Y(XZ+ZX)] = \mathrm{Tr}[Y\mathcal{L}_X(Z)]$ and $\mathrm{Tr}[\mathcal{L}_X(Y)Y] = 2\,\mathrm{Tr}[XY^2] \geq 2\underline{\lambda}(X)\,\mathrm{Tr}[Y^2]$ since $X - \underline{\lambda}(X)I_d \in \mathcal{S}_d^+$. The self-adjoint property is then clear for $\mathcal{L}_{X,a}$, and the positive definiteness comes from Lemma B.1 and the continuity of the eigenvalues of $\mathcal{L}_{X,a}$ with respect to a. □

Lemma B.3. *For $X \in \mathcal{S}_d^{+,*}$, $Y \in \mathcal{M}_d$, $\bar{\mathcal{L}}_X(Y) = \mathcal{L}_X^{-1}(YX + XY^\top)X$ is self-adjoint and positive. The linear application $\bar{\mathcal{L}}_{X,a}(Y) = \mathcal{L}_X^{-1}(YX + XY^\top)X - a\,\mathrm{Tr}[Y]I_d$ is also positive for $a < 1/\mathrm{Tr}[X^{-1}]$, and there is a positive $c_{X,a} > 0$ such that*

$$\mathrm{Tr}[\bar{\mathcal{L}}_X(Y)^\top Y] \geq c_{X,a}\,\mathrm{Tr}[(\mathcal{L}_X^{-1}(YX + XY^\top))^2].$$

Proof. Since $\mathcal{L}_X^{-1}$ is self-adjoint, we have for $Z \in \mathcal{M}_d$

$$\begin{aligned}\mathrm{Tr}[\bar{\mathcal{L}}_X(Y)^\top Z] &= \mathrm{Tr}[\mathcal{L}_X^{-1}(YX + XY^\top)ZX] = \frac{1}{2}\,\mathrm{Tr}[\mathcal{L}_X^{-1}(YX + XY^\top)(ZX + XZ^\top)] \\ &= \frac{1}{2}\,\mathrm{Tr}[(YX + XY^\top)\mathcal{L}_X^{-1}(ZX + XZ^\top)] = \mathrm{Tr}[Y^\top \bar{\mathcal{L}}_X(Z)].\end{aligned}$$

Similarly, $\mathrm{Tr}[\bar{\mathcal{L}}_{X,a}(Y)^\top Z] = \mathrm{Tr}[\bar{\mathcal{L}}_X(Y)^\top Z] - a\,\mathrm{Tr}[Y]\,\mathrm{Tr}[Z] = \mathrm{Tr}[Y^\top \bar{\mathcal{L}}_{X,a}(Z)]$. Besides, we notice that

$$\begin{aligned}\mathrm{Tr}[\bar{\mathcal{L}}_{X,a}(Y)^\top Y] &= \frac{1}{2}\left(\mathrm{Tr}[(YX + XY^\top)\mathcal{L}_X^{-1}(YX + XY^\top)] - 2a\,\mathrm{Tr}[Y]^2\right) \\ &= \frac{1}{2}\left(\mathrm{Tr}[\mathcal{L}_{X,a}(\mathcal{L}_X^{-1}(YX + XY^\top))\mathcal{L}_X^{-1}(YX + XY^\top)]\right)\end{aligned}$$

by Lemma B.1. This gives the claim since $\mathcal{L}_{X,a}$ is positive definite by Lemma B.2. □

The following lemma gives the Laplace transform of the matrix Normal distribution.

Lemma B.4. *Let $C \in \mathcal{S}_d^{+,*}$ and $\mathcal{C}[C] \in (\mathbb{R}^d)^{\otimes 4}$ defined by*

$$\mathcal{C}[C]_{i,j,k,l} = \delta_{ik}C_{j,l} + \delta_{il}C_{j,k} + \delta_{jk}C_{i,l} + \delta_{jl}C_{i,k}. \tag{71}$$

We introduce the $\mathcal{M}_d$-valued random variables $\tilde{\mathbf{G}}$ and $\mathbf{G} \sim \mathcal{N}(0, \mathcal{C}[C])$ of which components are Normal random variables with mean 0 such that

$$\forall i,j,k,l \in \{1,\dots,d\}, \quad \mathbb{E}[\tilde{\mathbf{G}}_{i,j}\tilde{\mathbf{G}}_{k,l}] = \delta_{ik}\delta_{jl}, \quad \mathbb{E}[\mathbf{G}_{i,j}\mathbf{G}_{k,l}] = \mathcal{C}[C]_{i,j,k,l}. \tag{72}$$

We have the following results.

1. *For all $c \in \mathcal{S}_d$, $\mathbb{E}\big[\exp(-\mathrm{Tr}[c\mathbf{G}])\big] = \exp(2\,\mathrm{Tr}[c^2 C])$.*
2. *For $\tilde{C} \in \mathcal{M}_d$ such that $\tilde{C}\tilde{C}^\top = C$, $\tilde{C}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top \tilde{C}^\top$ and $\mathbf{G}$ have the same law.*
3. *Let $X \in \mathcal{S}_d^{+,*}$. For $c \in \mathcal{S}_d$, $\mathbb{E}[\exp(-\mathrm{Tr}[c\mathcal{L}_X^{-1}(\sqrt{X}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top\sqrt{X})])] = \mathbb{E}[\exp(\mathrm{Tr}[c\mathcal{L}_X^{-1}(c)])$*

Proof. We focus on the first point. For all $c \in \mathcal{S}_d$, we have

$$\mathbb{E}\big[\exp(-\operatorname{Tr}[c\mathbf{G}])\big] = \mathbb{E}\big[\exp(-\sum_{1\leqslant i,j\leqslant d} c_{i,j}\mathbf{G}_{i,j})\big].$$

Moreover, $\sum_{1\leqslant i,j\leqslant d} c_{i,j}\mathbf{G}_{i,j}$ is a Normal random variable and its variance is given by

$$\mathbb{E}\big[\big(\sum_{1\leqslant i,j\leqslant d} c_{i,j}\mathbf{G}_{i,j}\big)^2\big] = \sum_{1\leqslant i,j,k,l\leqslant d} c_{i,j}c_{k,l}\mathcal{C}[C]_{i,j,k,l} = 4\operatorname{Tr}[c^2C].$$

It follows from the moment generating function of the Normal distribution that

$$\mathbb{E}\big[\exp(-\operatorname{Tr}[c\mathbf{G}])\big] = \exp(2\operatorname{Tr}[c^2C]).$$

To prove the second point it is sufficient to notice that $\operatorname{Tr}[c(\tilde{C}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top\tilde{C}^\top)] = \operatorname{Tr}[2c\tilde{C}\tilde{\mathbf{G}}]$ and

$$\mathbb{E}\big[\big(\sum_{1\leqslant i,j\leqslant d} (c\tilde{C})_{i,j}\tilde{\mathbf{G}}_{i,j}\big)^2\big] = \sum_{1\leqslant i,j,k,l\leqslant d} (c\tilde{C})_{i,j}(c\tilde{C})_{k,l}\delta_{ik}\delta_{jl} = \sum_{1\leqslant i,j\leqslant d} (c\tilde{C})^2_{i,j} = \operatorname{Tr}[c\tilde{C}\tilde{C}^\top c].$$

For the third point, we set $Z = \mathcal{L}_X^{-1}(\sqrt{X}\tilde{\mathbf{G}}+\tilde{\mathbf{G}}^\top\sqrt{X})$ and have $XZ+ZX = \sqrt{X}\tilde{\mathbf{G}}+\tilde{\mathbf{G}}^\top\sqrt{X}$. We also introduce $\tilde{c} = \mathcal{L}_X^{-1}(c)$ and have $\tilde{c}X + X\tilde{c} = c$. Thus, we obtain

$$\operatorname{Tr}[cZ] = \operatorname{Tr}[(\tilde{c}X + X\tilde{c})Z] = \operatorname{Tr}[\tilde{c}(\sqrt{X}\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^\top\sqrt{X})]$$

and therefore $\mathbb{E}[\exp(-\operatorname{Tr}[cZ])] = \exp(2\operatorname{Tr}[\tilde{c}^2X]) = \exp(\operatorname{Tr}[\tilde{c}(\tilde{c}X + X\tilde{c})]) = \exp(\operatorname{Tr}[\tilde{c}c])$. □

C Some asymptotic behaviour of Wishart processes

Lemma C.1. *Let $X \sim WIS_d(x,\alpha,b,I_d)$ with $b \in \mathcal{S}_d$, $x \in \mathcal{S}_d^+$ and $\alpha \geqslant d-1$. Then X_T converges in law when $T \to +\infty$ if and only if $-b \in \mathcal{S}_d^{+,*}$. In this case, X_T converges in law to $WIS_d(0,\alpha,0,\sqrt{-b^{-1}};1/2)$.*

Let $X \sim WIS_d(x,\alpha,b,I_d)$ with $b \in \mathcal{M}_d$, $x \in \mathcal{S}_d^+$ and $\alpha \geqslant d-1$. If $-(b+b^\top) \in \mathcal{S}_d^{+,*}$, $q_\infty := \int_0^\infty e^{sb}e^{sb^\top}ds$ is well defined and X_T converges in law to $WIS_d(0,\alpha,0,\sqrt{2q_\infty};1/2)$.

Proof. Let us first consider the case $-b \in \mathcal{S}_d^{+,*}$. From Proposition 4 in [2], we have for $v \in \mathcal{S}_d^+$,

$$\begin{aligned}\mathbb{E}[\exp(-\operatorname{Tr}[vX_T])] &= \frac{\exp\left(\operatorname{Tr}\left[-v\left(I_d + 2\left(\int_0^T e^{2bs}ds\right)v\right)^{-1}e^{Tb}xe^{Tb}\right]\right)}{\det\left[I_d + 2\left(\int_0^T e^{2bs}ds\right)v\right]^{\alpha/2}} \\ &\underset{T\to+\infty}{\to} \frac{1}{\det\left[I_d - b^{-1}v\right]^{\alpha/2}},\end{aligned}$$

which is the Laplace transform of $WIS_d(0,\alpha,0,\sqrt{-b^{-1}};1/2)$. Now, let us consider $-b \notin \mathcal{S}_d^{+,*}$. Then, there exists an eigenvector $v \in \mathbb{R}^d \setminus \{0\}$ such that $bv = \lambda v$ with $\lambda \geqslant 0$. Then, we have $\frac{d}{dt}\mathbb{E}[v^\top X_tv] = \alpha v^\top v + 2\lambda\mathbb{E}[v^\top X_tv]$, and therefore $\mathbb{E}[v^\top X_Tv] \underset{T\to+\infty}{\to} +\infty$.

In the case $b \in \mathcal{M}_d$ with $-(b+b^\top) \in \mathcal{S}_d^{+,*}$, we know that the norm of e^{bs} decays exponentially to 0 as $s \to +\infty$, see e.g. Problem 11.3.6 in Golub and Van Loan [18]. Using again Proposition 4 in [2], we get that $\mathbb{E}[\exp(-\operatorname{Tr}[vX_T])] \underset{T\to+\infty}{\to} \frac{1}{\det[I_d+2q_\infty v]^{\alpha/2}}$. $\square$

Lemma C.2. • *Assume $\alpha > d+1$ and $b=0$. Then, $\frac{Q_T^{-1}}{d\log(T)} \underset{T\to+\infty}{\to} \frac{1}{\alpha-(d+1)}$ a.s. Besides, $\frac{Z_T}{\log(T)}$ converges almost surely to d, and we have*

$$\forall \mu > 0, \sup_{T\geqslant 2} \mathbb{E}\left[\exp\left(\frac{\mu}{\sqrt{\log(T)}} N_T\right)\right] < \infty. \tag{73}$$

• *Assume $\alpha = d+1$ and $b=0$. Then, as $T\to+\infty$, $\left(\frac{2}{d\log(T)}\right)^2 Q_T^{-1}$ converges in law to $\tau_1 = \inf\{t \geqslant 0, B_t = 1\}$, where B is a Brownian motion. Besides, $\frac{Z_T}{\log(T)} = \frac{2N_T}{\log(T)}$ converges in probability to d, and we have*

$$\forall \mu > 0, \sup_{T\geqslant 2} \mathbb{E}\left[\exp\left(\frac{\mu}{\log(T)} N_T\right)\right] < \infty. \tag{74}$$

We mention that the results on the convergence for Q_T are given in Donati-Martin et al. [13]. However, their proofs is in a working paper by the same authors that we have not been able to find. For this reason, we present here an autonomous proof.

Proof. We first consider the case $\alpha > d+1$. We have $dX_t = \alpha I_d dt + \sqrt{X_t}dW_t + dW_t^\top\sqrt{X_t}$ and thus

$$d(e^{-t}X_{e^t-1}) = [\alpha I_d - e^{-t}X_{e^t-1}]dt + \sqrt{e^{-t}X_{e^t-1}}d\tilde{W}_t + d\tilde{W}_t^\top\sqrt{e^{-t}X_{e^t-1}},$$

with $d\tilde{W}_t = e^{-t/2}d(W_{e^t-1})$. We observe that $\tilde{W}$ is a matrix Brownian motion, which gives $Y \sim WIS_d(x,\alpha,-I_d/2,I_d)$, where $Y_t = e^{-t}X_{e^t-1}$ for $t \geqslant 0$. Using equation (25) to the process Y, we get

$$\frac{1}{t}\log\left(\frac{\det[Y_t]}{\det[Y_0]}\right) = (\alpha-1-d)\frac{1}{t}\int_0^t \operatorname{Tr}[Y_s^{-1}]ds - d + \frac{2}{t}\int_0^t \operatorname{Tr}[\sqrt{Y_s^{-1}}d\tilde{W}_s]. \tag{75}$$

Since Y is ergodic and $\langle \int_0^t \operatorname{Tr}[\sqrt{Y_s^{-1}}d\tilde{W}_s]\rangle = \int_0^t \operatorname{Tr}[Y_s^{-1}]ds$, we get that the left hand side converges in probability to zero and the right hand side converges a.s. to $(\alpha-1-d)\mathbb{E}[\operatorname{Tr}[Y_\infty^{-1}]]-d$, where $Y_\infty \sim WIS_d(0,\alpha,0,\sqrt{2}I_d;1/2)$ is the stationary law of Y. Therefore, $\frac{1}{t}\log\left(\frac{\det[Y_t]}{\det[Y_0]}\right)$ converges a.s. to zero. Since $\frac{1}{t}\log\left(\frac{\det[Y_t]}{\det[Y_0]}\right) = \frac{1}{t}\log\left(\frac{\det[e^{-t}X_{e^t-1}]}{\det[x]}\right) = \frac{1}{t}\log\left(\frac{\det[X_{e^t-1}]}{\det[x]}\right) - d$, we get that $\frac{Z_T}{\log(T)} = \frac{1}{\log(T)}\log\left(\frac{\det[X_T]}{\det[x]}\right)$ converges a.s. to d when $T\to+\infty$.

Now, we use (25) taken at time $T = e^t - 1$ and Dubins-Schwarz theorem: there is a Brownian motion β such that for all $t \geqslant 0$,

$$\frac{\alpha-(1+d)}{Q_{e^t-1}t} + \frac{2\beta_{Q_{e^t-1}^{-1}}}{t} = \frac{1}{t}\log\left(\frac{\det[X_{e^t-1}]}{\det[x]}\right).$$

This gives that $\frac{\alpha-(1+d)}{Q_{e^t-1}t} \underset{t\to+\infty}{\to} d$ a.s., and therefore $\frac{Q_T^{-1}}{d\log(T)} \underset{T\to+\infty}{\to} \frac{1}{\alpha-(d+1)}$, a.s.

It remains to prove (73). From (25), we have $N_T = \frac{Z_T}{2} - \frac{\alpha-1-d}{2}Q_T^{-1} \leqslant \frac{Z_T}{2}$ and thus $\mathbb{E}\left[\exp\left(\frac{\mu}{\sqrt{\log(T)}}N_T\right)\right] \leqslant \mathbb{E}\left[\left(\frac{\det[X_T]}{\det[x]}\right)^{\frac{\mu}{2\sqrt{\log(T)}}}\right] < \infty$, since the moments of X are bounded. Again we set $t = \log(T+1)$, and for $\Lambda \in [0,1]$, we have from (75)

$$\begin{aligned}N_T &= \int_0^T \mathrm{Tr}[\sqrt{X_s^{-1}}dW_s] = \int_0^t \mathrm{Tr}[\sqrt{Y_s^{-1}}d\tilde{W}_s]\\ &= \Lambda \int_0^t \mathrm{Tr}[\sqrt{Y_s^{-1}}d\tilde{W}_s] + (1-\Lambda)\left(\frac{1}{2}\log\left(\frac{\det[Y_t]}{\det[x]}\right) + \frac{d}{2}t - \frac{\alpha-1-d}{2}\int_0^t \mathrm{Tr}[Y_s^{-1}]ds\right).\end{aligned}$$

By Cauchy-Schwarz inequality, we get

$$\begin{aligned}&\mathbb{E}\left[\exp\left(\frac{\mu}{\sqrt{\log(T+1)}}N_T\right)\right]\\ \leq& e^{\frac{\mu d(1-\Lambda)}{2}\sqrt{t}}\mathbb{E}^{\frac{1}{2}}\left[\left(\frac{\det[Y_t]}{\det[x]}\right)^{(1-\Lambda)\frac{\mu}{\sqrt{t}}}\right]\\ &\times \mathbb{E}^{\frac{1}{2}}\left[\exp\left(\frac{2\mu\Lambda}{\sqrt{t}}\int_0^t \mathrm{Tr}[\sqrt{Y_s^{-1}}d\tilde{W}_s] - \mu(1-\Lambda)\frac{\alpha-1-d}{\sqrt{t}}\int_0^t \mathrm{Tr}[Y_s^{-1}]ds\right)\right].\end{aligned}$$

We now take $\Lambda = \Lambda_t = \frac{1}{2\epsilon_t}\left(-1+\sqrt{1+4\epsilon_t}\right)$ with $\epsilon_t = \frac{2\mu}{(\alpha-1-d)\sqrt{t}}$ in order to obtain $\frac{1}{2}\left(\frac{2\mu\Lambda_t}{\sqrt{t}}\right)^2 = \mu(1-\Lambda_t)\frac{\alpha-1-d}{\sqrt{t}}$. We note that for t large enough, $\Lambda_t \in [0,1]$. Besides, we have $\Lambda_t \underset{t\to+\infty}{=} 1 - \epsilon_t + o(1/t)$, so that $\sqrt{t}(1-\Lambda_t)$ converges to $\frac{2\mu}{\alpha-1-d}$. From Theorem 4.1 in [29], the second expectation is then equal to 1, while the first one is bounded since Y is ergodic. This yields to (73).

We now consider the case $\alpha = d+1$. We set again $t = \log(1+T)$ and have $T = e^t - 1$. Thus,

$$Z_T = \log\left(\frac{\det[X_T]}{\det[x]}\right) = \log\left(\frac{\det[e^t Y_t]}{\det[x]}\right) = \log\left(\frac{\det[Y_t]}{\det[x]}\right) + dt.$$

Again, Y_t converges in law towards $WIS_d(0,\alpha,0,\sqrt{2}I_d;1/2)$. Therefore, the ergodic theorem gives that $\frac{1}{t}\log\left(\frac{\det[Y_t]}{\det[x]}\right)$ converges in probability to 0, which yields to the convergence in probability of $\frac{Z_T}{\log(T)}$ to d. We now turn to the convergence of $\left(\frac{2}{d\log(T)}\right)^2 Q_T^{-1}$. We know from Theorem 4.1 in Mayerhofer [29] that for $T > 0$ and $\lambda \geqslant 0$,

$$\mathbb{E}\left[\exp\left(\frac{2\lambda}{d\log(1+T)}N_T - \frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right)\right] = 1.$$

From (25), we have $N_T = Z_T/2$ and we write

$$\begin{aligned}1 =&\mathbb{E}\left[\exp\left(\lambda - \frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right)\right]\\ &+\mathbb{E}\left[\exp\left(-\frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right)\left(\exp\left(\frac{2\lambda}{d\log(1+T)}N_T\right) - \exp(\lambda)\right)\right]\end{aligned}$$

We now observe that $\exp\left(-\frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right) \leqslant 1$ and that

$$\mathbb{E}\left[\exp\left(\frac{2\lambda}{d\log(1+T)}N_T\right)\right] = \mathbb{E}\left[\left(\frac{\det[X_T]}{\det[x]}\right)^{\frac{\lambda}{d\log(1+T)}}\right] = e^{\lambda}\mathbb{E}\left[\left(\frac{\det[Y_t]}{\det[x]}\right)^{\frac{\lambda}{dt}}\right].$$

Since Y has bounded moments and is stationary, $\sup_{t\geqslant 1}\mathbb{E}\left[\left(\frac{\det[Y_t]}{\det[x]}\right)^{\frac{\lambda}{dt}}\right]<\infty$. This gives the uniform integrability (74) and that

$$\mathbb{E}\left[\exp\left(-\frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right)\left(\exp\left(\frac{2\lambda}{d\log(1+T)}N_T\right)-\exp(\lambda)\right)\right]\underset{T\to+\infty}{\to}0.$$

Therefore, $\lim_{T\to+\infty}\mathbb{E}\left[\exp\left(\lambda-\frac{(2\lambda)^2}{2d^2\log(1+T)^2}Q_T^{-1}\right)\right]=1$, which gives the desired convergence in law. □

Acknowledgements. The authors would like to thank Arnaud Gloter (University of Evry) and Marina Kleptsyna (University of Le Mans) for helpful discussions, and the two anonymous referees for their fruitful comments.

References

[1] A. Ahdida and A. Alfonsi. Exact and high-order discretization schemes for Wishart processes and their affine extensions. *Ann. Appl. Probab.*, 23(3):1025–1073, 2013.

[2] A. Ahdida, A. Alfonsi, and E. Palidda. Smile with the Gaussian term structure model. *Working paper*, 2014.

[3] A. Alfonsi. *Affine diffusions and related processes: simulation, theory and applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan, 2015.

[4] M. Ben Alaya and A. Kebaier. Parameter estimation for the square-root diffusions: ergodic and nonergodic cases. *Stoch. Models*, 28(4):609–634, 2012.

[5] M. Ben Alaya and A. Kebaier. Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stoch. Anal. Appl.*, 31(4):552–573, 2013.

[6] M. Bru. *Thèse 3ème cycle. Résistance d'Escherichia coli aux antibiotiques: Sensibilités des analyses en composantes principales aux perturbations Browniennes et simulation.* PhD thesis, Université Paris Nord, 1987.

[7] M. Bru. Wishart processes. *J. Theoret. Probab.*, 4(4):725–751, 1991.

[8] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab.*, 21(2):397–463, 2011.

[9] J. Da Fonseca, M. Grasselli, and F. Ielpo. Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. *Stud. Nonlinear Dyn. Econom.*, 18(3):253–289, 2014.

[10] J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10:151–180, 2008.

[11] L. Dieci and T. Eirola. Positive definiteness in the numerical solution of Riccati differential equations. *Numer. Math.*, 67(3):303–313, 1994.

[12] G. Dohnal. On estimating the diffusion coefficient. *J. Appl. Probab.*, 24(1):105–114, 1987.

[13] C. Donati-Martin, Y. Doumerc, H. Matsumoto, and M. Yor. Some properties of the Wishart processes and a matrix extension of the Hartman-Watson laws. *Publ. Res. Inst. Math. Sci.*, 40(4):1385–1412, 2004.

[14] E. Fournié and D. Talay. Application de la statistique des diffusions à un modèle de taux d'interêt. *Finance*, 12(2):79–111, 1991.

[15] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 29(1):119–151, 1993.

[16] A. Gnoatto. The Wishart short rate model. *Int. J. Theor. Appl. Finance*, 15(8):1250056, 24, 2012.

[17] A. Gnoatto and M. Grasselli. The explicit Laplace transform for the Wishart process. *J. Appl. Probab.*, 51(3):640–656, 2014.

[18] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[19] C. Gourieroux and R. Sufana. Derivative pricing with Wishart multivariate stochastic volatility. *J. Bus. Econom. Statist.*, 28(3):438–451, 2010.

[20] C. Gourieroux and R. Sufana. Discrete time Wishart term structure models. *J. Econom. Dynam. Control*, 35(6):815–824, 2011.

[21] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[22] P. Jeganathan. Some aspects of asymptotic theory with applications to time series models. *Econometric Theory*, 11(5):818–887, 1995. Trending multiple time series (New Haven, CT, 1993).

[23] Y. A. Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London, Ltd., London, 2004.

[24] L. Le Cam. Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. california Publ. Statist.*, 3:37–98, 1960.

[25] L. Le Cam and G. L. Yang. *Asymptotics in statistics*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 2000. Some basic concepts.

[26] J. J. Levin. On the matrix Riccati equation. *Proc. Amer. Math. Soc.*, 10:519–524, 1959.

[27] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes. I,II*. Applications of Mathematics (New York). Springer-Verlag, Berlin, expanded edition, 2001. Applications, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[28] H. Luschgy. Local asymptotic mixed normality for semimartingale experiments. *Probab. Theory Related Fields*, 92(2):151–176, 1992.

[29] E. Mayerhofer. Wishart Processes and Wishart Distributions: An Affine Processes Point of View. *ArXiv e-prints*, Jan. 2012.

[30] B. Mond and J. E. Pečarić. On matrix convexity of the Moore-Penrose inverse. *Internat. J. Math. Math. Sci.*, 19(4):707–710, 1996.

[31] L. Overbeck. Estimation for continuous branching processes. *Scand. J. Statist.*, 25(1):111–126, 1998.

[32] G. Pagès. Sur quelques algorithmes récursifs pour les probabilités numériques. *ESAIM Probab. Statist.*, 5:141–170 (electronic), 2001.

[33] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[34] T. H. Rydberg. A note on the existence of unique equivalent martingale measures in a markovian setting. *Finance and Stochastics*, 1(3):251–257, 1997.

[35] D. W. Stroock. *Probability theory, an analytic view.* Cambridge University Press, Cambridge, 1993.